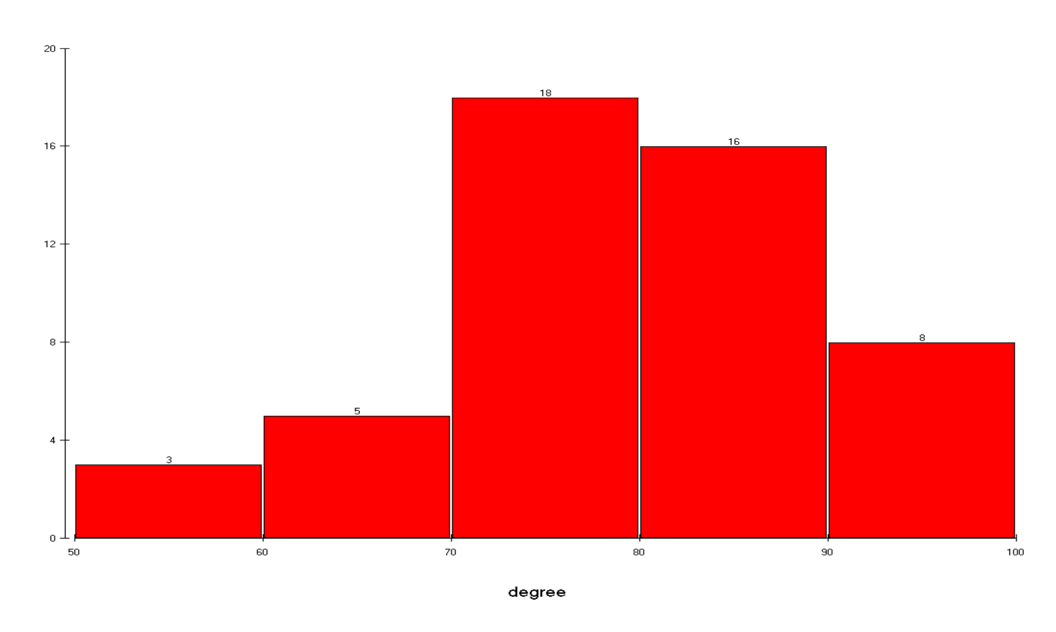
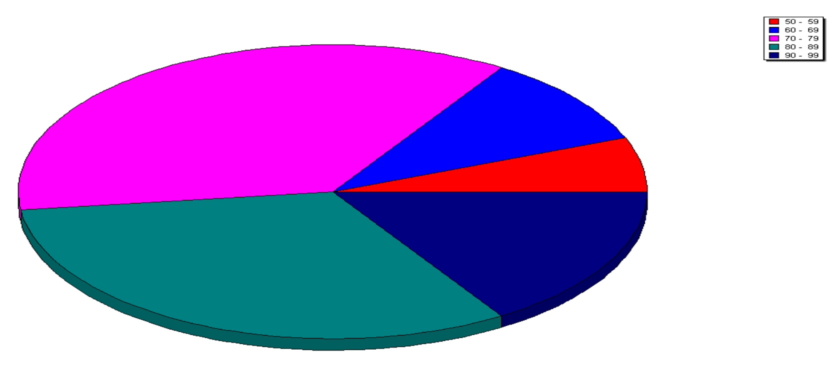
مبادئ الإحصاء والإحتمالات





صلاح علي مبخوت

المحتويات

المقدمة

Introduction

نشأ علم الإحصاء في العصور الوسطى لاهتمام الدول بتعداد أفراد المجتمع من حيث

(i) تكوين جيش قوي .

(ii) حصر ثروات الأفراد حتى تتمكن من فرض الضرائب.

(iii) بيانات عن المواليد والوفيات والإنتاج والاستهلاك.

وبذلك نشأت الحاجة إلى تنظيم هذه البيانات وتلخيصها ووضعها في صورة جداول أو رسم بياني حتى يسهل الرجوع إليها والاستفادة منها بأسرع وقت ممكن . وقد أطلق على هذه الطرق علم الدولة أو ″علم الملوك ″ انطلاقاً من كلمة ″status″ وتعني الدولة باللاتينية ثم اشتقت منها كلمة ″statistics″ أي علم الإحصاء لاسيما بعد تطور علم الاحتمالات في القرنين السابع عشر والثامن عشر بفضل جهود العلماء باسكال ″Pascal″ وبرنولي ″ Bernoulli″ و دمواڤر ″ Demoiver″ ولابلاس ″ Laplace″ وجاوس ″ Gauss″ .

بعد التطور في الإحصاء أصبحت الحاجة ملحة إلى تحليل البيانات التي تم جمعها كالتنبؤ بعدد السكان بعد فتره زمنية بناءً على التعداد الموجود أو التنبؤ بالإنتاج والاستهلاك . وقد ساعد على ذلك تطور علم الاحتمالات الذي كان له دور كبير في تحليل البيانات واتخاذ القرار المناسب . وقد امتد التطبيق الإحصائي إلى مجالات العلوم الأخرى كالطب والزراعة والفيزياء وفي القرن العشرين تطور علم الإحصاء بشكل كبير بواسطة الحاسوب.

الباب الأول

تعريف علم الإحصاء

يعرف علم الإحصاء بأنه ذلك الفرع من العلوم الذي يختص بالطرق العلمية لجمع البيانات وتنظيمها وتلخيصها وعرضها وتحليلها وذلك للوصول إلى نتائج مقبولة وقرارات سليمة وينقسم علم الإحصاء إلى قسمين أساسيين هما :

(i) **الإحصاء الوصفي** ″descriptive″ deductive statistics

وهو طرق تنظيم المعلومات وتلخيصها - الغرض من التنظيم هو المساعدة على فهم المعلومات – والطرق الوصفية تحتوي على توزيعات تكرارية (الجداول) ورسوم بيانية وطرق حساب مقاييس النزعة المركزية (المتوسطات) ومقاييس التشتت ومختلف القياسات الأخرى.

(ii) **الإحصاء الاستدلالي** inductive statistics

هو الوسائل العلمية التي تجري لدراسة نمو المجتمع(المعالِم) بناءً على المعلومات التي تم الحصول عليها من العينة المأخوذة منه وفق الطرق الإحصائية المعلومة .

المجتمع الإحصائي (population)

يعرف المجتمع بأنه مجموعة ذات خصائص مشتركة من الأفراد محل الدراسة ويقسم إلى محدود وغير محدود . في بعض الأحيان يكون من الصعب ملاحظة بيانات المجتمع ككل لما يكلف ذلك من جهد ومال أو قد يكون مستحيلاً . للتغلب على ذلك يمكن اختيار جزء من المجتمع يسمى بالعينة .

العينة الإحصائية (Sample)

العينة الإحصائية جزء من المجتمع تختار بحث تمثل جميع صفات المجتمع . إذ قد تكون الحاجة ضرورية لأخذ العينة بدلاً من دراسة المجتمع كله . مثل أخذ عينة من دم شخص لفحصها حيث إننا لا نستطيع فحص كل دم الشخص حتى لا يتوفى . وكذلك قد يؤدى دراسة المجتمع كله إلى فقدان عناصره أو إتلافها وهنا يجب أخذ عينة صغيره فمثلاً عند فحص سلامة كمية البيض يجب اخذ عينة منها ونقوم بكسرها لنرى ما إذا كان البيض سليماً أم لا . وتفيد المعلومات المتوفرة من العينات في التنبؤ عن معلومات ومؤشرات عن سلوك المجتمع كله .

المَعلَم و الإحصاءة والمتغير (parameter)

المعلم (parameter) : شيء يميز المجتمع كله وذلك مثل متوسط الدخل الشهري للأسرة في دولة معينة.

الإحصاءة (statistic) : شيء يميز العينة مثل متوسط الدخل الشهري لعينة مكونة من 100 أسرة من مجتمعٍ ما .

المتغير : هو مقدار له خصائص رقمية (كمية) وغير رقمية (وصفية) تتغير قيمه من عنصر إلى آخر من عناصر المجتمع أو العينة . فمثلاً متغير الطول أو الوزن كمية في حين متغير الجنس أو اللون متغيرات وصفية .

تنظيم البيانات وعرضها جدولياً

مثال:البيانات الآتية تمثل درجة 50 طالباً في إحدى المواد :

51	95	70	74	73	90	71	74	90	67
91	72	83	89	50	80	72	84	85	69
62	82	87	76	91	76	87	75	78	79
71	96	81	88	64	82	93	57	86	70
80	81	75	85	74	90	83	66	77	91

تنظيم البيانات الإحصائية فيما يسمى جدول التوزيع التكراري frequency distribution تفرغ منه البيانات آنفة الذكر كالآتي :

نختار عدد الفئات =5 ثم نوجد طول الفئة $\dfrac{\text{المدى للتكرارات}}{\text{عدد الفئات}}$

$$L = \frac{R}{\text{n}} = \frac{96-51}{5} = 9 \approx 10$$

جدول تفريغ وتوزيع الدرجات

التكرار f_i (عدد الطلاب)	العلامات	الصفات الفئات
3	111	50 – 59
5	~~1111~~	60 – 69
18	~~1111~~ ~~1111~~ ~~1111~~ 111	70 – 79
16	~~1111~~ ~~1111~~ ~~1111~~ 1	80 – 89
8	~~1111~~ 111	90 – 99
50		المجموع

جدول التوزيع التكراري

التكرار	**حدود الفئات**
3	50 – 59
5	60 – 69
18	70 – 79
16	80 – 89
8	90 – 99
50	**المجموع**

و يمكن أن يكتب أفقياً لتوفير حيز الكتابة ، كالآتي

المجموع	90 – 99	80 – 89	70 – 79	60 – 69	50 – 59	**الفئات**
50	8	16	18	5	3	**التكرار**

جدول التوزيع التكراري النسبي

التكرار النسبي	حدود الفئات
0.06	50 – 59
0.10	60 – 69
0.36	70 – 79
0.32	80 – 89
0.16	90 – 99
1	المجموع

جدول التوزيع التكرارى المئوى

حدود الفئات	التكرار المئوي
50 – 59	6
60 – 69	10
70 – 79	36
80 – 89	32
90 – 99	16
المجموع	100

الحدود الحقيقية (الفعلية) للفئات :

البيانات الإحصائية عادة تكون مكتوبة مقربة مثلاً لأقرب وحدة قياس أو لأقرب نصف وحدة قياس . فإذا كانت البيانات مقربة لأرقام صحيحة فإننا نطرح من الحد الأدنى المقرب للفئة 0.5 لنحصل على الحد الأدنى الحقيقي ونضيف 0.5 إلى الحد الأعلى المقرب لنحصل على الحد الأعلى الحقيقي للفئة . أما إذا كانت البيانات لأقرب رقم عشري فإننا نطرح أو نضيف 0.05 وهكذا .

جدول التوزيع التكراري بالحدود الحقيقية (الفعلية) للفئات

التكرار	**حدود الفئات**
3	49.5 – 59.5
5	59.5 – 69.5
18	69.5 – 79.5
16	79.5 – 89.5
8	89.5 – 99.5
50	**المجموع**

مركز الفئات : مركز الفئة = $\frac{\text{الحد الأعلى للفئة + الحد الأدنى للفئة}}{2}$

التكرار المئوي	التكرار النسبي	التكرار	مركز الفئات	الحدود الفعلية	حدود الفئات
6	0.06	3	54.5	49.5	50 – 59
10	0.10	5	64.5	59.5	60 – 69
36	0.36	18	74.5	69.5	70 – 79
32	0.32	16	84.5	79.5	80 – 89
16	0.16	8	94.5	89.5	90 – 99
100	1.00	50			المجموع

الجدول التكراري المتجمع الصاعد ("less than" cumulative freq)

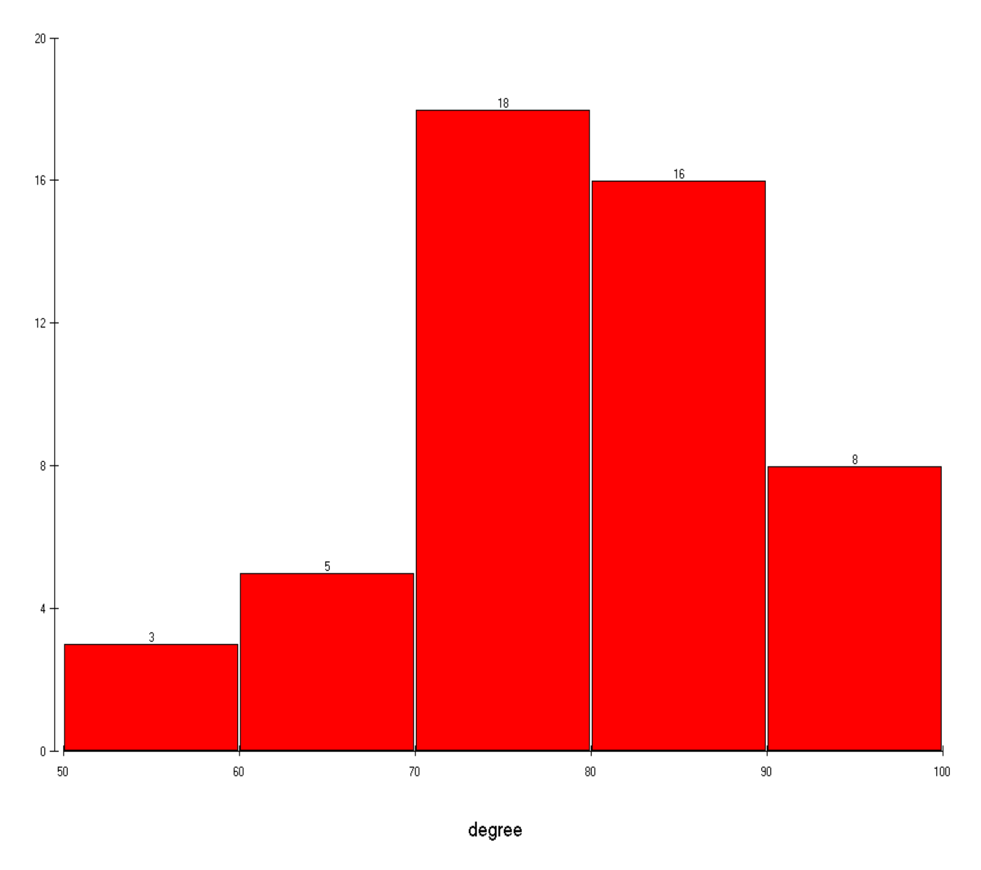
التكرار المتجمع الصاعد	حدود الفئات
0	أصغر من 49.5
3	أصغر من 59.5
8	أصغر من 69.5
26	أصغر من 79.5
42	أصغر من 89.5
50	أصغر من 99.5

فمثلاً عدد الطلاب الحاصلين على درجات أقل من 79 درجة : 3+8+18=26

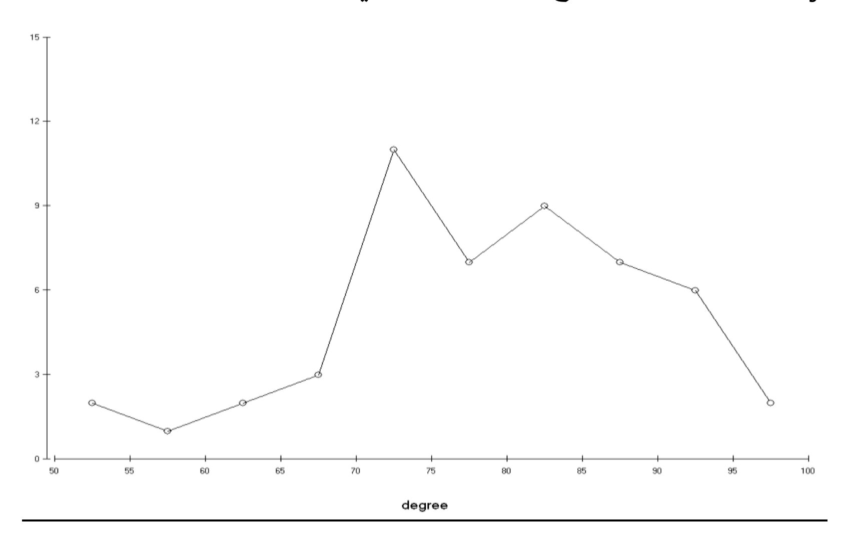
الجدول التكراري المتجمع الهابط ("more than" cumulative frequency)

التكرار المتجمع الهابط	حدود الفئات
50	أكبر من 49.5
47	أكبر من 59.5
42	أكبر من 69.5
24	أكبر من 79.5
8	أكبر من 89.5
0	أكبر من 99.5

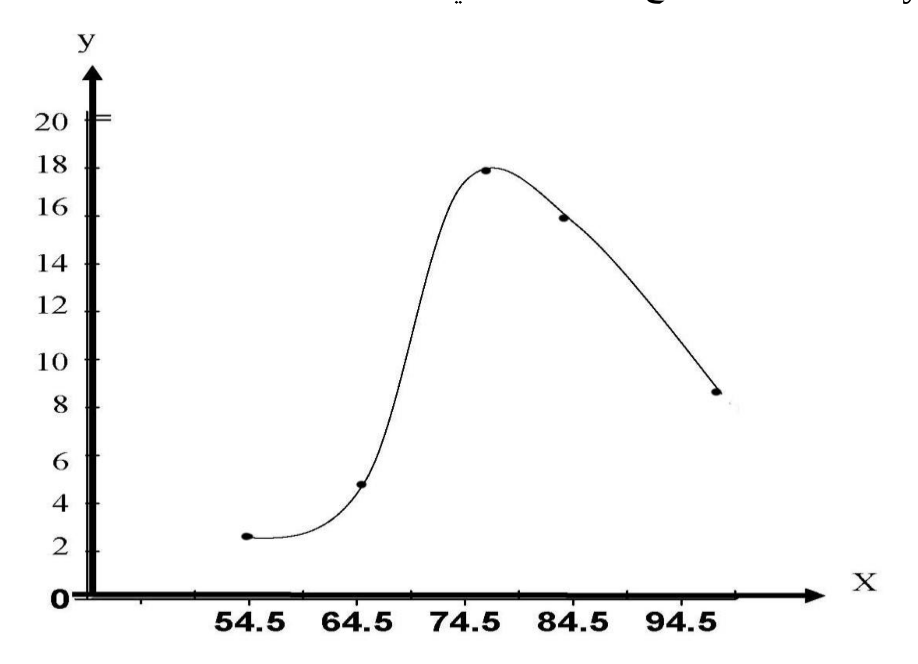
مدرج التكرار: و يرسم كعلاقة بين مركز الفئات x والتكرار y على هيئة مستطيلات على النحو الآتي :



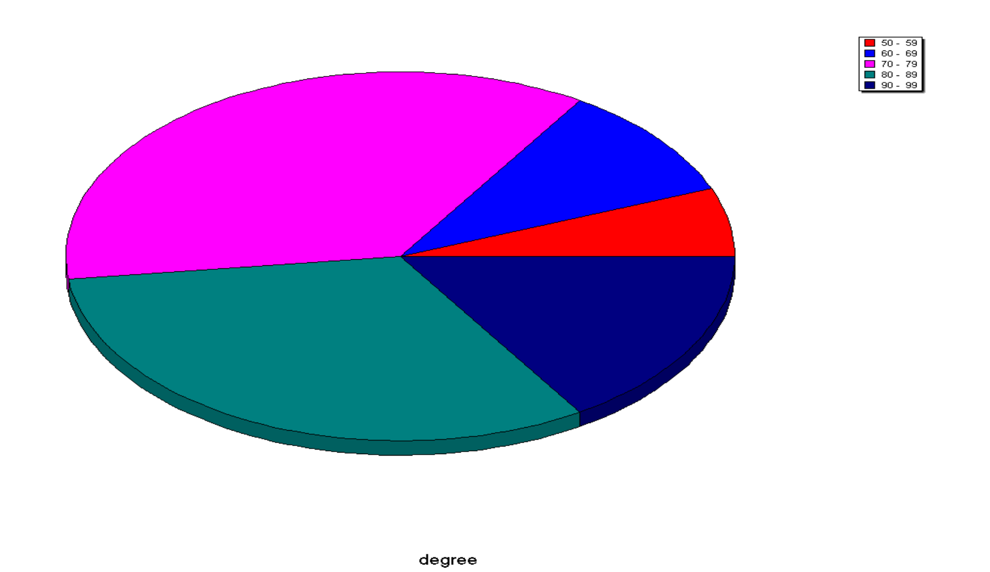
<u>مضلع التكرار</u> : ويرسم كعلاقة بين مركز الفئات - محور x - والتكرار - محور y - وعلى هيئة أضلاع على النحو الآتي :



منحنى التكرار : ويرسم كعلاقة بين مركز الفئات - محور x - والتكرار - محور y - وعلى هيئة أضلاع على النحو الآتي :



الشكل الدائري للتوزيع Pie Diagram



تمرين

نظم البيانات التالية و لخصها جدولياً ومن ثم اعرضها بيانياً :

10	11	12	16	13	14	16	22	27	32	35
39	40	28	29	30	22	24	26	32	33	35
26	37	32	31	29	28	29	40	40	28	22
24	23	27	26	29	31	32	34	33	31	30
31	29	27	29	28	32	34	39	38	18	21
23	19	17	18	20						

الباب الثاني

مقاييس النزعة المركزية (المتوسطات)

من المعروف عادةً أن الرسوم البيانية غير دقيقة بغرض الحصول على بعض خصائص المجتمع الإحصائي محل الدراسة . لذلك يجب أن يكون لدينا مقاييس عددية تصف لنا هذه البيانات وسوف نستعرض في هذا الباب نوع مهم من المقاييس الإحصائية وهو ما يسمى بمقاييس النزعة المركزية أو المتوسطات.

ويمكن تعريف المتوسطات بأنها القيمة النموذجية الممثلة لمجموعة من البيانات وحيث أن القيمة النموذجية تميل إلى الوقوع في المركز لذلك فإنه يمكن أن تسمى المتوسطات بمقاييس النزعة المركزية.

وبعض المتوسطات الأكثر شيوعاً هي الوسط الحسابي (أو المتوسط) mean والوسيط median والمنوال mode وهنالك أيضا الوسط الهندسي والوسط التوافقي وذلك في حالة البيانات المباشرة والمئوية.

تعريف رمز المجموع $\sum$

إذا كان لدينا مجموعة من المشاهدات

$$x_1, x_2, x_3, \ldots\ldots, x_n$$

فإن حاصل جمع هذه المشاهدات يعبر عنه بالمجموع

$$\sum_{i=1}^{n} x_i = x_1 + \ldots\ldots + x_2$$

الوسط الحسابي أو المتوسط (mean)

يعتبر من أهم مقاييس الموضع (النزعة المركزية) والأكثر استخداما في الإحصاء . إذا كان لديننا مجموعة من المشاهدات للمتغير X

وهي $x_1, x_2, x_3, \ldots\ldots, x_n$ فإن الوسط الحسابي يساوى حاصل جمع المشاهدات مقسوماً على عددها ويرمز له بالرمز $\overline{x}$.

$$\overline{x} = \frac{x_1 + x_2 + x_3 + \ldots + x_n}{n} = \frac{1}{n}\sum_{i=1}^{n} x_i$$

<u>مثال 1</u>

إذا كانت درجات 5 طلاب في إحدى المواد هي 60,72,40,80,63 ، احسب المتوسط ؟

الحل : $\overline{x} = \frac{1}{n}\sum_{i=1}^{n} x_i = \frac{1}{5}\sum_{i=1}^{5} x_i$

$$= \frac{1}{5}(60 + 72 + 40 + 80 + 63) = \frac{1}{5}(315) = 63$$

<u>الوسط الحسابي للبيانات المبوبة</u>

إذا كان لدينا عدد K من الفئات ذات المراكز $x_1, x_2, x_3, \ldots\ldots, x_k$ لها تكرارات $f_1, f_2, \ldots, f_k$ على الترتيب فإن الوسط الحسابي يعطي بالعلاقة الآتية

$$\overline{x} = \frac{x_1 f_1 + x_2 f_2 + x_3 f_3 + \ldots\ldots + x_k f_k}{f_1 + f_2 + \ldots + f_n}$$

$$= \frac{\sum_{i=1}^{k} x_i f_i}{\sum_{i=1}^{k} f_i} = \frac{1}{n}\sum x_i f_i$$

حيث $n = \sum_{i=1}^{k} f_i$

مثال 2

احسب متوسط أعمار الطلاب $\overline{X}$ للبيانات الآتية

فئات الأعمار	5 – 6	7 – 8	9 – 10	11 – 12	13 – 14
عدد الطلاب	2	5	8	4	1

الحل : لسهولة الحل نكوّن الجدول الآتي

الفئات	مركز الفئات X	التكرار f	$f.X$
5 – 6	5.5	2	11
7 – 8	7.5	5	37.5
9 – 10	9.5	8	76.0
11 – 12	11.5	4	46.0
13 – 14	13.5	1	13.5
المجموع $\sum$		$n = \sum f_i = 20$	$\sum f_i x_i = 184$

بالتعويض في قانون الوسط الحسابي نحصل على

$$\overline{x} = \frac{1}{n} \sum x_i f_i = \frac{1}{20}(184) = 9.2$$

خواص المتوسط

$$\sum_{i=1}^{n} \left(x_i - \overline{x} \right) = \sum_{i=1}^{n} d_i = 0 \qquad (i)$$

أي أن المجموع الجبري لانحرافات القيم عن وسطها الحسابي يساوي صفر

$$\overline{(ax \pm b)} = a\overline{x} \pm b \qquad (ii)$$

(iii) مميزات المتوسط :

(١) مقياس سهل حسابه ويخضع للعمليات الجبرية بسهولة .

(٢) يأخذ في الاعتبار جميع القيم محل الدراسة .

(٣) أكثر المقاييس فهماً في الإحصاء .

(iv) عيوب المتوسط

(١) يتأثر بالقيم المتطرفة (الكبيرة جداً والصغيرة جداً)

(٢) يصعب حسابه في حالة البيانات الوصفية

الوسيط Median

عند ترتيب البيانات ترتيباً تصاعدياً – أو تنازلياً – فالوسيط يكون هو القيمة التي يقع 50% من البيانات قبلها في الترتيب و 50 % من البيانات بعدها في الترتيب. فإذا كان عدد البيانات فردياً يكون الوسيط هو المشاهدة التي في المنتصف وإذا كان عدد البيانات زوجياً فإن الوسيط هو متوسط المشاهدتين التين في المنتصف.

مثال 3

أوجد الوسيط لدرجات الطلاب الآتية

$$60,72,40,80,63$$

الحل :

نرتب البيانات تصاعدياً : $40, 60, \boxed{63}, 72, 80$

وبما أن عدد المشاهدات – البيانات – فردي فإن الوسيط هو المشاهدة التي تقع في المنتصف

$$\text{med} = 63$$

مثال 4

أوجد الوسيط لدرجات الطلاب الآتية :

$$72,60,72,40,80,63$$

الحل :

نرتب البيانات تصاعدياً : $40,60,\boxed{63},\boxed{72},72,80$

بما أن عدد البيانات زوجي فإن الوسيط هو $med = \frac{63+72}{2} = 67.5$

الوسيط من الجداول التكرارية

الفئة الوسطية هي الفئة التي يقع فيها الوسيط .

لإيجاد الوسيط حسابياً نتبع الخطوات الآتية :

(١) نكون الجدول المتجمع الصاعد (باستخدام الحدود الحقيقية)

(٢) نوجد رتبة الوسيط $\frac{n}{2}$ (سواء كانت n فردية أو زوجية)

(٣) نحدد مكان الوسيط بعد حساب $\frac{n}{2}$ بين التكرارات المتجمعة في الجدول المتجمع الصاعد و نضع خطاً أفقياً يمر داخل الفئة الوسيطية و يكون التكرار المتجمع السابق لهذا الخط هو f_1 والتكرار المتجمع اللاحق هو f_2 .

(٤) نحدد البداية الحقيقية للفئة الوسيطية و يرمز لها L .

(٥) نعيِّن طول الفئة الوسيطية h ويساوى الحد الأدنى للفئة التالية مطروحاً منه الحد الأدنى للفئة الوسيطية .

(٦) يعطي الوسيط بالعلاقة $med = L + \dfrac{\frac{n}{2} - f_1}{f_2 - f_1} h$

<u>البرهان</u>

حيث f_1 التكرار المتجمع الصاعد السابق للتكرار المتجمع الوسيطي .

f_2 التكرار المتجمع الصاعد اللاحق للتكرار المتجمع الوسيطي .

A

X

$\frac{n}{2} - f_1$

f_1

$\frac{n}{2}$

$f_2 - f_1$

L

f_2

بإجراء التناسب بين الأطوال والتكرارات في الشكل السابق و بوضع

$A = L, L = h$ نحصل على

$$\frac{x}{h} = \frac{\frac{n}{2} - f_1}{f_2 - f_1}$$

$$x = \frac{\frac{n}{2} - f_1}{f_2 - f_1} h$$

$$med = L + x$$

$$= L + \frac{\frac{n}{2} - f_1}{f_2 - f_1} .h$$

مثال 5

احسب الوسيط لأعمار الطلاب في مثال 2

فئات الأعمار	5 – 6	7 – 8	9 – 10	11 – 12	13 – 14
عدد الطلاب	2	5	8	4	1

الحل : نكون الجدول التكراري المتجمع الصاعد كالآتي

فئات التكرارات المتجمعة الصاعدة	التكرارات المتجمعة الصاعدة
أصغر من 4.5	0
أصغر من 6.5	2
أصغر من 8.5	7 f_1
L	
أصغر من 10.5	15 f_2
أصغر من 12.5	19
أصغر من 14.5	20

نحسب $\frac{n}{2}$ وهي تساوى $\frac{n}{2}=\frac{20}{2}=10$

ونلاحظ أن 10 تقع بين 15 و 7

ونضع خطاً أفقيا يمثل تكرار الوسيط المتجمع 10 وعليه فيكون

$$h = 10.5 - 8.5 = 2$$

$$f_2 = 15$$

$$f_1 = 7$$

$$L = 8.5$$

وبتطبيق قانون الوسيط نحصل على

$$med = L + \frac{\frac{n}{2} - f_1}{f_2 - f_1} h = 8.5 + \frac{10 - 7}{15 - 7} .2 = 8.5 + \frac{6}{8} = 9.25$$

المنوال Mode

هو القيمة الأكثر تكراراً في مجموعة البيانات . قد يكون لمجموعة البيانات منوال واحد لذلك تسمى وحيده المنوال أو يكون لها أكثر من منوال وتسمى متعددة المنوال أو لا يكون لمجموعة البيانات منوال فتسمى عديمة المنوال .

مثال 6

أحسب المنوال من البيانات الآتية :

2.6.9.4.10.6

الحل

يوجد لهذه البيانات منوال واحد وهو القيمة 6 لأنها تكررت 3 مرات أكثر من غيرها .

مثال 7

احسب المنوال من البيانات الآتية

4.2.7.9.4.7.10.7

الحل:

نجد أن القيمة 7 تكررت 3 مرات و 4 تكررت مرتين وعليه فإن المنوال هو 7

مثال 8

احسب المنوال من البيانات الآتية

4.9.8.12.11.7.15

الحل:

لا يوجد في هذه البيانات أي قيمة تكررت أكثر من مرة وعليه فإنه لا يوجد منوال لهذه البيانات .

المنوال من جداول التوزيع التكرارية

١ - نوجد أكبر تكرار f و عليه يمكن إيجاد التكرار السابق له f_1 و اللاحق f_2

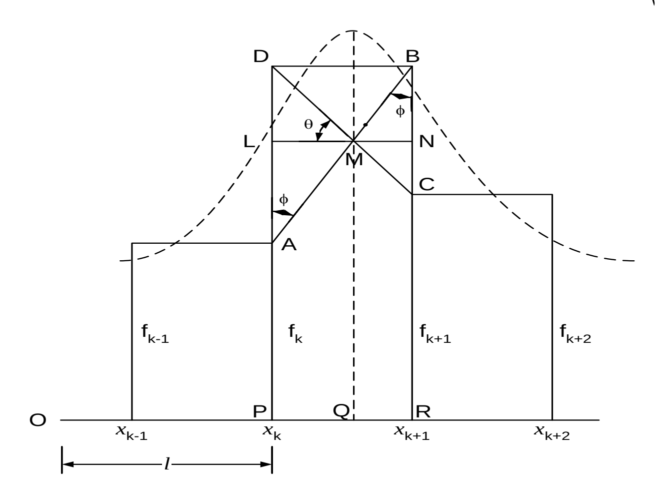
٢ - نأخذ بداية الفئة المنوالية و يرمز له بالرمز L و هي الفئة التي تقابل أكبر تكرار f .

٣ - نحدد طول الفئة المنوالية h و هو يساوي الفرق بين بداية الفئة المنوالية و بداية الفئة التالية لها و نطبق القاعدة الآتية :

$$Mod = L + \frac{f - f_1}{2f - f_1 - f_2}.h$$

نبرهن قاعدة المنوال السابقة على النحو الآتي

من الرسم



$$\tan\phi=\frac{LM}{AL}=\frac{MN}{NB}. \quad \text{و} \quad \tan\theta=\frac{LD}{LM}=\frac{NC}{MN}$$

لذلك

$$\text{و} \quad \frac{LM}{MN}=\frac{LD}{NC}=\frac{AL}{NB}=\frac{AL+LD}{NB+NC}=\frac{AD}{BC}$$

$$\frac{LM}{LN-LM}=\frac{PD-AP}{BR-CR}$$

$$\frac{LM}{h-LM}=\frac{f_k-f_{k-1}}{f_k-f_{k+1}},$$

$$LM=\frac{h(f_k-f_{k-1})}{(f_k-f_{k+1})+(f_k-f_{k-1})}=\frac{h(f_k-f_{k-1})}{2f_k-f_{k+1}-f_{k-1}}.$$

،

حيث h طول الفئة . إذن

<u>مثال9</u>

أوجد المنوال حسابياً لأعمار الطلاب في المثال 2

فئات العمر	5 – 6	7 – 8	9 – 10	11 – 12	13 – 14
عدد الطلاب	2	5	8	4	1

الحل:

من الجدول نجد أن

$$f = 8 \quad f_2 = 4 \quad f_1 = 5$$

وكذلك $h = 10.5 - 8.5 = 2 \quad , L = 8.5$

بالتعويض في قانون المنوال السابق نحصل على

$$= 8.5 + \frac{8-5}{16-5-4} .2$$

<u>مثال 10</u>

الأعمار	6	7	8	9	10	11
الطلاب	4	2	7	3	2	2

مراكز الفئات	التكرار	$f_i x_i$	المتجمع الصاعد
0			0 أقل من 6
6	4	24	4 أقل من 7
[7]	2	14	6 أقل من 8
8	7	56	13 أقل من 9
9	3	27	16أقل من 10
10	2	20	18أقل من 11
11	2	22	20أقل من 12
المجموع	20	163	

$$(i)\ \bar{x} = \frac{\sum f_i x_i}{\sum f_i} = \frac{163}{20}$$

$$(ii)\ \frac{n}{2} = 10, L = 7, f_1 = 6, f_2 = 13$$

$$Med = L + \frac{\frac{n}{2} - f_1}{f_2 - f_1}.h = 7 + \frac{10-6}{13-6}.1 = 7 + \frac{4}{7}$$

$$(iii) f = 7, f_1 = 2 f_2 = 3, L = 8, h = 9 - 8 = 1$$

$$Mod = L + \frac{f - f_1}{2f - f_2 - f_1}.h = 8 + \frac{7-2}{14-3-2}.1$$

$$= 8 + \frac{5}{9}$$

مثال 11:

احسب المنوال و الوسيط للبيانات الآتية:

x_i	0	1	2	3	4
f_i	20	10	8	5	3

ملاحظة : من تعريف المنوال عندما تجد في البيانات فئة تتمتع بتكرار أعلى من تكرار أي فئة أخرى و يتناقص التكرار باطراد في الفئات السابقة لها أو اللاحقة و نقول أن هذه الفئة هي الفئة المنوالية و نعتبر مركزها منوالاً للبيان الإحصائي

$$\therefore \text{mod} = 0$$

x_i	f_i	$f_i x_i$	المتجمع الصاعد
0	20	0	0 أقل من 0
1	10	10	20 أقل من 1
2	8	16	30 أقل من 2
3	5	15	38 أقل من 3
4	3	12	43 أقل من 4
10	46	53	46 أقل من 5

$$\frac{n}{2} = \frac{46}{2} = 23, L = 1, f_1 = 20, f_2 = 30, h = 2 - 1 = 1$$

$$Med = L + \frac{\frac{n}{2} - f_1}{f_2 - f_1}.h = 1 + \frac{23 - 20}{14 - 3 - 2}.1 = 1.3$$

الوسط الهندسي(Geometric Mean)

الوسط الهندسي لمجموعة من القيم $x_1,\ldots,x_n$ ، هو الجذر النوني لحاصل ضرب

هذه القيم ، كالآتي : $G.M = \sqrt[n]{x_1 x_2 \ldots x_n}$

يمتاز الوسط الهندسي عن الوسط الحسابي بأنه أقل تأثراً بالقيم المتطرفة في البيانات ، لأنه معلوم رياضياً بأن الوسط الهندسي لمجموعة من القيم أقل من وسطه الحسابي $G.M \leq \bar{x}$. وعادةً يحسب الوسط الهندسي باستخدام اللوغريتمات كالآتي:

$$Log\ G.M = \frac{1}{n}\left(\sum_{l=1}^{n} Log\ x_i\right)$$

مثال12:

أحسب الوسط الهندسي و الوسط الحسابي للبيانات

3,5,6,6,7,10,12

الحل:

$$G.M = \sqrt[7]{3.5.6.6.7.10.12}$$

$$LogG.M = \frac{1}{7}(Log3 + Log5 + Log6 + Log6 + Log7 + Log10 + Log12)$$

$$= \frac{1}{7}(0.4771 + 0.6990 + 0.7782 + 0.7782 + 0.8451 + 1 + 1.0729) = 0.8081$$

$$Log\ G.M = 0.8081$$

$$\Rightarrow G.M = 6.43$$

$$\bar{x} = \frac{1}{7}(3+5+6+6+7+10+12) = 7$$

من الواضح أن $G.M \leq \bar{x}$

الوسط الهندسي في حالة الجداول التكرارية

في هذه الحالة يحسب الوسط الهندسي للفئات التي عددها K و مراكزها هي $x_1, x_2, \ldots, x_k$ و التي يقابلها بالترتيب التكرارات $f_1, \ldots, f_k$ من القانون الآتي :

$$G.M = \sqrt[n]{x_1^{f_1} \ldots x_k^{f_k}}$$

الوسط التوافقي (Harmonic mean)

يستخدم الوسط التوافقي عندما يكون مقلوب المتغير له دلالة ، كأن يعين نسبة بين متغيرين مثل السرعة بالنسبة للزمن . و الوسط التوافقي H لمجموعة n من القيم $x_1, \ldots, x_n$ هو مقلوب الوسط الحسابي لمقلوبات هذه القيم أي أن

$$H = \frac{1}{\frac{1}{n}\sum_{i=1}^{n}\frac{1}{x_i}}$$

$$\frac{1}{H} = \frac{1}{n}\sum_{i=1}^{n}\frac{1}{x_i}$$

مثال13:

أحسب الوسط التوافقي H للبيانات الآتية

3,5,6,6,7,10,12

$$\frac{1}{H} = \frac{1}{7}\left(\frac{1}{3}+\frac{1}{5}+\frac{1}{6}+\frac{1}{6}+\frac{1}{7}+\frac{1}{10}+\frac{1}{12}\right) = \frac{501}{2940}$$

$$\therefore H = \frac{2940}{501} = 5.87$$

الوسط التوافقي للبيانات المبوبة

إذا كانت مراكز الفئات $x_1,\ldots,x_k$ و تكرارها $f_1,\ldots,f_k$ على الترتيب فإنه

$$\frac{1}{H}=\frac{1}{n}\sum_{i=1}^{n}\frac{f_i}{x_i}$$

حيث $n=\sum_{i=1}^{k}f_i$

الرُبَيْعات

إذا رتبت عينة من البيانات حسب قيمها تصاعدياً أو تنازلياً فإن القراءة التي في المنتصف و التي تقسم العينة إلى مجموعتين متساويتين في العدد هي الوسيط كما سبق تعريفه و بتعميم الفكرة و تقسيم البيانات بعد ترتيبها إلى أربعة أجزاء متساوية فإن نقاط التقسيم يرمز لها Q_1,Q_2,Q_3 حيث Q_1 الرُبيع الأول و Q_2 يسمى الرُبيع الثاني و Q_3 يسمى الرُبيع الثالث (و الرُبيع الثاني Q_2 يساوي الوسيط).

العُشَيْريات:

بصورة مشابهة نجد العشيريات $D_1, D_2,\ldots,D_{10}$, حيث أن D_1 العشير الأول و يسبقه $\frac{1}{10}$ من القراءات ، D_2 و يسبقه $\frac{2}{10}$ من القراءات ، .... الخ

المئينات:

بصورة مشابهة نجد المئينات $P_1,P_2,\ldots,P_{99}$ حيث P_1 يسمى المئين الأول و يسبقه $\frac{1}{100}$ من القرارات ، ...الخ

و في حالة البيانات المبوبة يعطى قانون حساب العشيريات و المئينات كالوسيط

مع استبدال $\frac{n}{2}$ بالرقم $\frac{n}{10}$ للعشير الأول و $\frac{2n}{10}$ للعشير الثاني و هكذا دواليك . و

استبدال $\frac{n}{2}$ بالرقم $\frac{n}{100}$ للمئين الأول، $\frac{2n}{100}$ للمئين الثاني وهكذا دواليك .

مثال14:

أوجد كلاً من العشير الثاني و المئين التسعين لأعمار الطلاب

المجموع	13 – 14	11 – 12	9 – 10	7 – 8	5 – 6	فئات الأعمار
20	1	4	8	5	2	عدد الطلاب

الحل

التكرار المتجمع الصاعد	حدود الفئات	
0	أقل من 4.5	
2	أقل من 6.5	D_2
7	أقل من 8.5	
15	أقل من 10.5	P_{90}
19	أقل من 12.5	
20	أقل من 14.5	

لإيجاد العشير الثاني

$$D_2 = L + \frac{\frac{2n}{10} - f_1}{f_2 - f_1}.h = 6.5 + \frac{4-2}{7-2}.2 = 7.3$$

لإيجاد المئين التسعين

$$P_{90} = \overline{L} + \frac{\frac{90n}{100} - \overline{f}_1}{\overline{f}_2 - \overline{f}_1}.h = 10.5 + \frac{18-15}{29-15}.2 = 10.93$$

التمارين

١. فيما يلي أعمار مجموعة من الطلاب بإحدى المدارس

6,6,9,8,6,10,9,9,8,7,8,6,7,8,8,11,10,11,8,8

أ - أحسب المتوسط الحسابي

ب - احسب المنوال

ت - احسب الوسيط

٢. فيما يلي توزيع درجات 60 طالباً في أحد الاختبارات

فئة	40-44	45-49	50-54	55-59	60-64	65-69	70-74	75-79	80-84	85-89	90-94	95-99
التكرار	3	3	4	6	6	11	9	8	2	3	3	1

أحسب: أ - المتوسط ب - الوسيط ت - المنوال

٣. فيما يلي أطوال مجموعة من الطلاب في أحد المدارس

الأطوال	65	66	67	68	69	70	71	72	73	74
التكرار	3	3	4	6	6	11	9	8	2	3

احسب: أ - المتوسط ب - الوسيط ت - المنوال

٤. فيما يلي توزيع الأجر اليومي للعد من العمال بالريال في أحد المصانع

فئات الأجر	20-29	30-39	40-49	50-59	60-69	70-79	المجموع
عدد العمال	9	12	15	8	4	2	50

احسب: أ - الوسط الحسابي لأجور العمال ب - الوسيط والمنوال حسابياً ت - الوسط الهندسي والوسط التوافقي للأجور

٥. فيم يلي تصنيف لعدد أيام الغياب خلال فصل دراسي لشعبة تتضمن 46 طالباً

عدد أيام الغياب	5	1	2	3	4
عدد الطلاب	20	10	8	5	3

أحسب: أ - المتوسط ب - الوسيط ت - المنوال

الباب الثالث

مقاييس التشتت Measures of Dispersion

مقدمة

لنفترض أن لدينا ثلاث مجموعات مختلفة من الطلاب X,Y,Z درجاتهم كالآتي :

$$X = 59, 61, 62, 58, 60$$
$$Y = 50, 60, 66, 54, 70$$
$$Z = 19, 65, 46, 78, 72$$

نجدهم جميعاً يتفقون في المتوسط الحسابي $\overline{x} = 60$. في حين نجد درجات المجموعة الأولى متقاربة و متجانسة ، درجات المجموعة الثانية أقل تقارباً و تجانساً من المجموعة الأولى ، درجات المجموعة الثالثة أقل تقارباً من المجموعة الثانية . أي أن الثلاث مجموعات مختلفة التجانس رغم اتفاقها في المتوسط ، و من ثم فإن مقاييس النزعة المركزية غير كافية للمقارنة بين طبيعة البيانات الإحصائية . من هنا نشأت الحاجة إلى إيجاد مقاييس تقيس درجة تجانس (تقارب) أو تشتت (تباعد) مفردات البيانات بعضها عن بعض فيما يعرف بمقاييس التشتت و منها المدى ، الانحراف المتوسط ، التباين ، الانحراف المعياري ، معامل الاختلاف و مقاييس الالتواء و التفلطح .

المدى

يُعرَّف المدى ، في حالة البيانات المباشرة ، بأنه الفرق بين أصغر قراءة و أكبر قراءة و في حالة البيانات المبوبة فإن المدى يكون الفرق بين مركز الفئة العليا و مركز الفئة الدنيا . و لأن المدى يعتمد على قراءتين فقط فهو مقياس تقريبي لا يعتمد عليه .

نصف المدى الرُبيعي

إذا كانت لدينا مجموعة من البيانات عددها n قراءة فإن القراءات ترتب ترتيباً تصاعدياً وتقسم إلى أربعة أقسام متساوية، ويعطى نصف المدى الربيعي بالعلاقة

$$Q = \frac{Q_3 - Q_1}{2}$$

حيث Q_1 يسمى الربيع الأدنى ويسبقه $\frac{n}{4}$ من القراءات

و Q_3 يسمى الربيع الأعلى ويسبقه $\frac{3n}{4}$ من القراءات

مثال 1

أوجد نصف المدى الربيعي لأوزان مجموعة الطلاب الآتية

$$67, 65, 69, 58, 55, 71, 72, 70$$

الحل

نعيد ترتيب البيانات تصاعدياً، كالآتي

$$55, 58 \uparrow 65, 67, 69, 70 \uparrow 71, 72$$

ثم نوجد

$$Q_1 = \frac{58+65}{2} = 61.5$$

$$Q_3 = \frac{70+71}{2} = 70.5$$

$$Q = \frac{Q_3 - Q_1}{2}$$

$$= \frac{70.5 - 61.5}{2} = 4.5$$

مثال 2

أوجد نصف المدى الربيعي لأوزان مجموعة الطلاب الآتية

$$59, 67, 65, 69, 58, 50, 70, 72, 74$$

الحل

نعيد ترتيب البيانات تصاعدياً، كالآتي

$$50, 58, [59], 65, 67, 69, [70], 72, 74$$

ثم نوجد

$$Q_1 = 59$$
$$Q_3 = 70$$
$$Q = \frac{Q_3 - Q_1}{2}$$
$$= \frac{70 - 59}{2} = 5.5$$

نصف المدى الربيعي للبيانات المبوبة

يتم حساب نصف المدى الربيعي بصورة مشابهة لحساب الوسيط :

يحسب الربيع الأدنى Q_1 بوضع $\frac{n}{4}$ بدلاً من $\frac{n}{2}$ في قانون الوسيط .

يحسب الربيع الأعلى Q_2 بوضع $\frac{3n}{4}$ بدلاً من $\frac{n}{2}$ في قانون الوسيط .

وأخيراً يحسب نصف المدى الربيعي من القانون

$$Q = \frac{Q_3 - Q_1}{2}$$

$$Q_1 = L_1 + \frac{\frac{n}{4} - F_1}{F_2 - F_1} h$$

$$Q_3 = L_2 + \frac{\frac{3n}{4} - F_1^{`}}{F_2^{`} - F_1^{`}} h$$

مثال 3

أوجد نصف المدى الربيعي حسابياً لدرجات الطلاب الآتية

حدود الفئات	التكرار المئوي
40 -49	2
50 – 59	9
60 – 69	15
70 – 79	11
80 – 89	2
90 – 99	1

الحل

نوجد الجدول المتجمع الصاعد كما يلي

التكرارات المتجمعة الصاعدة	فئات التكرارات المتجمعة الصاعدة
0	أصغر من 39.5
$2,...F_1$	أصغر من $L_1 = 49.5$
$11,...F_2$	أصغر من 59.5
$26,...F_1^{`}$	أصغر من $L_2 = 69.5$
$37,...F_2^{`}$	أصغر من 79.5
39	أصغر من 89.5
40	أصغر من 99.5

لاحظ :

$$n = 40, \frac{n}{4} = 10, \frac{3n}{4} = 30, h = 10$$

$$Q_1 = L_1 + \frac{\frac{n}{4} - F_1}{F_2 - F_1} h = 49.5 + \frac{10-2}{11-2}.10 = 58.39$$

$$Q_3 = L_2 + \frac{\frac{3n}{4} - F_1^{`}}{F_2^{`} - F_1^{`}} h = 69.5 + \frac{30-26}{37-26}.10 = 73.14$$

$$Q = \frac{Q_3 - Q_1}{2} = \frac{73.14 - 58.39}{2} = 7.38$$

الانحراف المتوسط (Mean deviation)

يعرف الانحراف المتوسط بأنه متوسط الانحرافات المطلقة للقراءات عن متوسطها الحسابي $\bar{x}$ و يرمز له بالرمز M.D . أي أن

$$M.D = \frac{1}{n}\sum_{i=1}^{n}|x_i - \bar{x}|$$

و السبب في أخذ القيم المطلقة للانحرافات هو أن مجموع انحرافات القيم عن متوسطها يساوي صفرا .

في حالة البيانات المبوبة فإن الانحراف المتوسط يعطى بالعلاقة الآتية

$$M.D = \frac{1}{n}\sum_{i=1}^{n}f\ |x_i - \bar{x}|$$

مثال 4

أوجد الانحراف المتوسط لأعمار مجموعة الطلاب

$$6, 5, 7, 7, 8, 9, 9, 5$$

الحل : أولاً نوجد المتوسط $\bar{x} = \frac{56}{8}$. ثم نوجد جدول الحل كالآتي

x	$x - \bar{x}$	$\lvert x - \bar{x}\rvert$
6	-1	1
5	-2	2
7	0	0
7	0	0
8	1	1
9	2	2

9 5	2 -2	2 2
56	0	10

$$M.D = \frac{1}{n}\sum_{i=1}^{n}|x_i - \bar{x}| = \frac{10}{8} = 1.25$$

مثال 5

أوجد الانحراف المتوسط لدرجات الطلاب في مثال 3 السابق

الحل : أولاً نوجد المتوسط $\bar{x} = \frac{\sum fx}{\sum f} = \frac{2630}{40} = 65.75$

الفئات	x	f	fx	$x-\bar{x}$	$\|x-\bar{x}\|$	$f\|x-\bar{x}\|$
40 -49	44.5	2	89	−21.25	21.25	42.5
50 – 59 60 – 69 70 – 79 80 – 89 90 – 99 $\sum$	54.5 64.5 74.5 84.5 94.5	9 15 11 2 1 40	490.5 967.5 819.5 169.0 94.5 2630	−11.25 −1.25 8.75 18.75 28.75	11.25 1.25 8.75 18.75 28.75	101.25 18.75 96.25 37.5 28.75 325

$$M.D = \frac{1}{n}\sum_{i=1}^{n} f\ |x_i - \bar{x}| = \frac{325}{40} = 8.125$$

التباين و الانحراف المعياري

تباين المجتمع هو متوسط انحرافات القيم عن وسطها الحسابي و يرمز له بالرمز σ^2 و يقرأ سيجما تربيع . في حالة البيانات المباشرة ، إذا كان لدينا قراءات من مجتمع إحصائي عدد مفرداته N هي : $x_1, x_2, ..., x_N$ و متوسطها $\bar{x}$ فإن مربعات انحرافات هذه القيم عن $\bar{x}$ هي

$$(x_1 - \bar{x})^2, (x_2 - \bar{x})^2, ..., (x_N - \bar{x})^2$$

ومن ثم فإن التباين

$$\sigma^2 = \frac{(x_1 - \bar{x})^2 + (x_2 - \bar{x})^2 + ... + (x_N - \bar{x})^2}{N}$$

$$\sigma^2 = \frac{1}{N}\sum_{i=1}^{N}(x_i - \bar{x})^2$$

الانحراف المعياري للمجتمع هو الجذر التربيعي للتباين ، أي أن

$$\sigma = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(x_i - \bar{x})^2}$$

تباين العينة

بالنسبة لعينة حجمها n من مجتمع حجمه N فإن تباين العينة يرمز له s^2

$$s^2 = \frac{1}{n-1}\sum_{i=1}^{n}(x_i - \bar{x})^2$$

الانحراف المعياري للعينة يرمز له S

$$s = \sqrt{\frac{1}{n-1}\sum_{i=1}^{n}\left(x_i - \bar{x}\right)^2}$$

في حالة البيانات المباشرة فإن S يعطي تقدير أفضل للانحراف المعياري σ .

إذا كان حجم العينة كبير $(n \geq 30)$ فإن S^2 و σ^2 متساويتان تقريباً من الناحية العملية .

<u>مثال 6</u>

احسب الانحراف المعياري S لأعمار مجموعة من الطلاب الآتية

$$8, 9, 7, 6, 5$$

الحل

أولاً نوجد المتوسط $\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i = \frac{35}{5} = 7$

و من ثم نوجد جدول الحل كالآتي

x	$x - \bar{x}$	$\left(x - \bar{x}\right)^2$
8	1	1
9	2	4
7	0	0
6	1-	1
5	2-	2
$\sum$	0	10

$$s^2 = \frac{1}{n-1}\sum_{i=1}^{n}(x_i - \bar{x})^2 = \frac{10}{5-1} = 2.5$$
$$s = \sqrt{2.5} = 1.591$$

التباين و الانحراف المعياري للبيانات المبوبة

إذا كانت $x_1, x_2, ..., x_k$ مراكز فئات تكراراتها $f_1, f_2, ...f_k$ على التوالي فإن

$$s^2 = \frac{1}{n-1}\sum_{i=1}^{n} f\ (x_i - \bar{x})^2$$
$$s = \sqrt{\frac{1}{n-1}\sum_{i=1}^{n} f\ (x_i - \bar{x})^2}$$

مثال 7

أوجد الانحراف المعياري لدرجات الطلاب الآتية

الفئة	40−49	50−59	60−69	70−79	80−89	90−99
الدرجة	2	9	15	11	2	1

الفئات	x	f	fx	$x - \bar{x}$	$(x - \bar{x})^2$	$f(x - \bar{x})^2$
40 -49	44.5	2	89	−21.25	451.56	903.13
50 – 59	54.5	9	490.5	−11.25	126.56	1139.06
60 – 69	64.5	15	967.5	−1.25	1.56	23.06
70 – 79	74.5	11	819.5	8.75	76.56	842.19
80 – 89	84.5	2	169.0	18.75	351.56	703.13
90 – 99	94.5	1	94.5	28.75	826.56	826.56
$\sum$		40	2630			4437.5

لاحظ أن
$$\bar{x} = \frac{1}{n}\sum_{i=1}^{n} f_i x_i = \frac{2630}{40}$$

$$s^2 = \frac{1}{n-1}\sum_{i=1}^{n} f_i \left(x_i - \bar{x}\right)^2$$
$$= \frac{4437.5}{39} = 113.78$$
$$s = \sqrt{113.78} = 10.67$$

مثال 8

أثبت أن
$$s^2 = \frac{1}{n-1}\left[\sum_{i=1}^{n} x_i^2 - \frac{\left(\sum_{i=1}^{n} x_i\right)^2}{n}\right]$$

الحل

$$s^2 = \frac{1}{n-1}\sum_{i=1}^{n}\left(x_i - \bar{x}\right)^2$$
$$= \frac{1}{n-1}\sum(x^2 - 2x\bar{x} + \bar{x}^2)$$
$$= \frac{1}{n-1}\left(\sum x^2 - 2\bar{x}\sum x + \sum \bar{x}^2\right)$$
$$= \frac{1}{n-1}\left(\sum x^2 - 2\bar{x}.n\bar{x} + n\bar{x}^2\right)$$
$$= \frac{1}{n-1}\left(\sum x^2 - n\bar{x}^2\right)$$
$$= \frac{1}{n-1}\left(\sum x^2 - \frac{\left(\sum x\right)^2}{n}\right)$$

مثال 9

احسب الانحراف المعياري لأعمار الطلاب الآتية

$$8, 9, 7, 6, 5$$

الحل

	x	x^2
	8	64
	9	81
	7	49
	6	36
	5	25
$\sum$	35	255

$$s^2 = \frac{1}{n-1}\left[\sum_{i=1}^{n} x_i^2 - \frac{\left(\sum_{i=1}^{n} x_i\right)^2}{n}\right]$$

$$= \frac{1}{4}\left(255 - \frac{(35)^2}{5}\right) = \frac{1}{4}(255 - 245) = 2.5$$

$$s = \sqrt{2.5} = 1.581$$

تمارين

١ ـ احسب الانحراف المتوسط و الانحراف المعياري للبيانات الآتية

x	11–15	16–20	21–25	26–30
f	**12**	**14**	**13**	**11**

2 – احسب الانحراف المتوسط و الانحراف المعياري للبيانات الآتية

Class	**Frequency**
11–15	**12**
15–20	**14**
20–25	**13**
25–30	**11**

الباب الرابع

الارتباط Correlation و الانحدار Regression

مقدمة

تناولنا في الفصول السابقة طرق دراسة متغير واحد لأي ظاهرة مثل أوزان مجموعة من الطلاب أو أعمارهم أو الأجور لمجموعة من العمال و تلخيصها جدولياً و عرضها بيانياً . كذلك درسنا بعض المقاييس العددية التي تساعد على معرفة بعض خصائص التوزيعات التكرارية و منها المتوسطات و التشتت .
الآن سوف نتناول دراسة البيانات التي يكون لأفرادها متغيران يتغيران معاً في آنٍ واحد و ذلك لمعرفة نوع العلاقة التي تربط بينهما مثل دراسة العلاقة بين أوزان و أطوال مجموعة من الطلاب أو أعمار و درجات مجموع من الطلاب أو العلاقة بين الدخل و الإنفاق أو بين الأجور و الإنتاج . و من ثم نوجد مقاييس تقيس درجة هذه العلاقة . أيضاً سنتناول العلاقة بين المتغيرين $\left(X ,Y \right)$ و إمكانية التعبير عن تلك العلاقة بمعادلة رياضية تعيننا على التنبؤ بقيمة أحد المتغيرين إذا علمنا قيمة المتغير الآخر . سوف نكتفي بإيجاد مقاييس تقيس قوة الارتباط بين المتغيرين $\left(X ,Y \right)$ في حالة العلاقة الخطية فقط .

معامل الارتباط الخطي لبيرسون

يستخدم معامل الارتباط الخطي لبيرسون لقياس التغير الذي يطرأ على المتغير Y عندما تتغير قيم X و العكس . في حالة البيانات المباشرة ، إذا كان لدينا زوج المشاهدات الآتية من مجتمع : $\left(x_1, y_1\right), \left(x_2, y_2\right), ..., \left(x_3, y_3\right)$ فإن معامل الارتباط لبيرسون يرمز له r و يعطى من خلال العلاقة الآتية

$$r = \frac{1}{n}\frac{\sum_{i=1}^{n}(x_i - \bar{x})(y_i - \bar{y})}{\sigma_x \sigma_y} \quad ...(1)$$

σ_x , σ_y هما الانحراف المعياري للمتغيرين y و x على الترتيب . في حالة البيانات المأخوذة من عينة فإن العلاقة السابقة تصبح كالآتي

$$r = \frac{1}{n-1}\frac{\sum_{i=1}^{n}(x_i - \bar{x})(y_i - \bar{y})}{\sigma_x \sigma_y} \quad ...(2)$$

لاحظ : معامل الارتباط يساوي صفر عندما تكون الظاهرتان مستقلتين و يكون قوياً عندما يقترب معامل الارتباط من الواحد الصحيح و ضعيفاً عندما يقترب من الصفر .لإيجاد معامل ارتباط بيرسون حسابياً يفضل استخدام الصيغة الآتية

$$r = \frac{n\sum_{i=1}^{n} x_i y_i - \sum_{i=1}^{n} x_i \sum_{i=1}^{n} y_i}{\sqrt{\left(n\sum_{i=1}^{n} x_i^2 - \left(\sum_{i=1}^{n} x_i\right)^2\right)\left(n\sum_{i=1}^{n} y_i^2 - \left(\sum_{i=1}^{n} y_i\right)^2\right)}} \quad ...(3)$$

<u>مثال 1</u>

الجدول الآتي يبين درجات مجموعة مكونة من ثمانية طلاب في كلٍ من مادتي الإحصاء X و الرياضيات Y . هل توجد علاقة بين تحصيل الطلاب في المادتين ؟

X	Y
13	15
9	7
19	17
15	15
11	10
8	9
16	14
11	10

الحل: لتبسيط البيانات التي بالجدول نطرح مقداراً ثابتاً $a = 10$ من كل من قيم X و Y و من ثم نكون الجدول الآتي

X	Y	$x' = x - 10.$	$y' = y - 10.$	$x'y'$	x'^2	y'^2
13	15	3	5	15	9	25
9	7	1-	3-	3	1	9
19	17	9	7	63	81	49
15	15	5	5	25	25	25
11	10	1	0	0	1	0
8	9	2-	1-	2	4	1
16	14	6	4	24	36	16
11	10	1	0	0	1	0

نعوض عن قيم $x = x'$ و $y = y'$ في معادلة 3 لنحصل على معامل ارتباط بيرسون

$$r = \frac{8\times132-22\times17}{\sqrt{\left(8\times158-(22)^2\right)\left(8\times125-(17)^2\right)}}$$

$$r = \frac{690}{744.7} = 0.93 \approx 1$$

و هذا يعني وجود ارتباط قوي جداً بين درجات تحصيل الطلاب في المادتين .

<u>معامل الارتباط للرتب لسبيرمان Spearman</u>

معامل الارتباط الخطي لبيرسون الذي سبق الحديث عنه يقيس مقدار قوة الارتباط بين متغيرين وذلك في حالة البيانات الكمية فقط . في بعض الأحيان يكون مطلوباً إيجاد قوة الارتباط بين متغيرين على صورة بيانات وصفية مثل تقديرات الطلاب في مادتين مختلفتين معطاة بالأحرف A,B,C,D,E فإنه يصعب حساب معامل الارتباط بطريقة بيرسون ، لذلك نشأت الحاجة إلى إيجاد مقياس يعطي قوة الارتباط للبيانات الوصفية وهذا المقياس هو ما يسمى بمعامل ارتباط الرتب لسبيرمان و هو يعطي مقياس للارتباط في كل من البيانات الكمية و الوصفية التي لها صفة الترتيب مثل تقديرات الطلاب فإنه يمكن إعطاءها رتب من حيث كبر التقدير أو صغره . نلاحظ أن رتب المتغيرين (X, Y) تزيد و تنقص حسب زيادة و نقص قيمهما ، لذلك فإن حساب معامل الارتباط للرتب يقترب كثيراً من معامل الارتباط لبيرسون ولكن يمتاز عنه في السهولة و الدقة و خاصة عندما تكون أزواج القيم أقل من 30 و يعطى معامل ارتباط الرتب بالعلاقة الآتية

$$r_s = \frac{1-6\sum_{i=1}^{n} d_i}{n\left(n^2-1\right)}$$

حيث r_s معامل ارتباط الرتب لسبيرمان لعدد n من أزواج القيم (X,Y) و حيث $d_i = a_i - b_i$ حيث a_i رتب المتغير X و b_i رتب المتغير Y. ويمكن توضيح حساب الرتب كالآتي : المقصود بالرتب هنا هو إيجاد رتب القراءات (X,Y) مع بقاء كل قراءة مكانها و ذلك بأن نتصور ترتيب البيانات تصاعدياً أو تنازلياً ، ففي حالة الترتيب التصاعدي تأخذ رتبة أصغر قراءة الرقم 1 في حين تأخذ القراءة التي تليها الرقم 2 وهكذا . في حالة تساوي قيمتان نعطي لكل قيمة منهما رتبة تساوي الوسط الحسابي لرتبتيهما .

<u>مثال 2</u>

أوجد رتب X التي قيمها معطاة في الجدول الآتي

X	10	4	5	7	2

الحل

نتصور ترتيب قيم X تصاعدياً كما في الجدول الآتي

X	10	4	5	7	2
الرتبة	2	9	15	11	2

<u>مثال 3</u>

أوجد رتب التقديرات الآتية

$$B, C, B, E, D, D, A$$

الحل

إذا تصورنا E تأخذ الرتبة 1 ، D مكررة و من ثم تأخذ الرتبتين 2 و 3 و عليه كل قيمة للتقدير D تأخذ الرتبة $2.5 = \frac{2+3}{2}$ ، C تأخذ الرتبة 4 ، B مكررة و تأخذ الرتبة $5.5 = \frac{5+6}{2}$ وأخيراً A تأخذ الرتبة 7 كما في الجدول الآتي

X	A	D	D	E	B	C	B
الرتبة	7	2.5	2.5	1	5.5	4	5.5

مثال 4

أوجد معامل ارتباط الرتب لدرجات الطلاب في مادتي الإحصاء و الرياضيات:

إحصاء	13	9	19	15	8	16	11
رياضيات	15	7	17	10	9	14	10

الحل

نكون الجدول الآتي

إحصاء	رياضيات	a	b	d	d^2
13	15	5	6.5	−1.5	2.25
9	7	2	1	1	1
19	17	8	8	0	0
15	15	6	6.5	−0.5	2.25

11	1	3.5	3.5	0	0
8	9	1	2	–0.5	1
16	14	7	5	2	4
11	10	3.5	3.5	0	0
$\sum$					8.5

$$r_s = \frac{1-6\sum_{i=1}^{n} d_i}{n\left(n^2-1\right)} = 1-\frac{6\times 8.5}{8\left(64-1\right)} = 1-\frac{51}{504} = 0.9$$

معامل الاقتران Coefficient of contingency

يستخدم معامل الاقتران لقياس قوة الارتباط بين ظاهرتين ، كل ظاهرة منهما ذات صفتين فقط و سوف يرمز له بالرمز c.c ، مثلاً دراسة علاقة قوة الارتباط بين التدخين و التعليم . الجدول الآتي يبين التكرار للصفات

	يدخن	لايدخن
متعلم	A	B
غير متعلم	C	D

فيكون معامل الاقتران كالآتي

$$c.c = \frac{AD-BC}{AD+BC}$$

مثال 5

عند دراسة علاقة التدخين بالتعليم في إحدى المؤسسات اختيرت عينة مكونة من 17 شخص وكانت النتائج موضحة كالآتي

لا يدخن	يدخن	
5	5	متعلم
4	3	غير متعلم

احسب معامل الاقتران بين التدخين و التعليم ؟

الحل

$$c.c=\frac{AD-BC}{AD+BC}=\frac{5\times4-3\times5}{5\times4+3\times5}=\frac{5}{35}=0.14$$

معامل التوافق

إذا كانت بيانات الظاهرتين موزعة على أكثر من نوعين (أي أن الجدول يحتوي على أكثر من أربع خانات) فإن معامل الاقتران السابق لا يصلح في هذه الحالة ونستخدم مقياساً آخر هو معامل التوافق . لحساب معامل التوافق نفرض أن الظاهرة X لها r صفة والظاهرة Y لها s صفة . نوضح جدول الاقتران بين الظاهرتين كما يلي

x/y	y_1	y_2	.	.	y_s	$\sum$
x_1	f_{11}	f_{12}	.	.	f_{1s}	$f_{1.}$
x_2	f_{21}	f_{22}	.	.	f_{2s}	$f_{2.}$
.	.	.	.	.	.	.
.	.	.	.	.	.	.
x_r	f_{r1}	f_{r2}	.	.	f_{rs}	$f_{r.}$
$\sum$	$f_{.1}$	$f_{.2}$			$f_{.s}$	$f_{..}$

و من ثم نحسب معامل التوافق كالآتي

$$c.c = \frac{\sqrt{B-1}}{B}$$

$$B = \frac{(f_{11})^2}{f_{.1} \times f_{1.}} + \frac{(f_{12})^2}{f_{.2} \times f_{2.}} + \ldots + \frac{(f_{rs})^2}{f_{.s} \times f_{r.}}$$

<u>مثال 6</u>

عند دراسة العلاقة بين الرائحة و لون الزهور لعينة مكونة 30 زهرة توصلنا للنتائج الآتية

x/y	yes	no	$\sum$
$yellow$	6	4	10
$white$	7	2	9
red	6	5	11
$\sum$	19	11	30

أحسب معامل التوافق بين اللون X و الرائحة Y للزهور ؟

الحل

$$B=\frac{(6)^2}{19\times10}+\frac{(7)^2}{19\times9}+\frac{(6)^2}{19\times11}+\frac{(4)^2}{11\times10}+\frac{(2)^2}{11\times9}+\frac{(5)^2}{11\times11}$$

$$=0.19+0.29+0.17+0.15+0.04+0.21=1.05$$

$$c.c=\frac{\sqrt{B-1}}{B}=\frac{\sqrt{1.05-1}}{1.05}=\frac{\sqrt{0.05}}{1.05}=0.22$$

نلاحظ أن مقدار قوة الارتباط ضعيفة .

خط الانحدار linear regression

معادلة خط الانحدار هي التنبؤ بقيمة المتغير التابع لقيمة محددة من قيم المتغير المستقل .

١ ـ معادلة خط انحدار Y على X ، و تعطى بالعلاقة $y=mx+c$ حيث

$$m=\frac{n\sum xy-\sum x\sum y}{n\sum x^2-\left(\sum x\right)^2}$$

$$c=\frac{\sum y}{n}-m\frac{\sum x}{n}$$

٢ ـ معادلة خط انحدار X على Y ، و تعطى بالعلاقة $x=my+c$ حيث

$$m=\frac{n\sum xy-\sum x\sum y}{n\sum y^2-\left(\sum y\right)^2}$$

$$c=\frac{\sum x}{n}-m\frac{\sum y}{n}$$

مثال 7

أوجد معادلة خط انحدار Y على X و معادلة خط انحدار X على Y حيث X درجات الإحصاء و Y درجات الرياضيات ، في الجدول الآتي

إحصاء	13	9	19	15	8	16	11
رياضيات	15	7	17	10	9	14	10

الحل : نكون الجدول الآتي

x	y	xy	x^2	y^2
13	15	195	169	225
9	7	63	81	49
19	7	323	361	289
15	15	225	225	225
11	10	110	121	100
8	9	72	64	81
16	14	224	256	196
11	10	110	121	100
$\sum = 102$	97	1322	1398	1265

١ -معادلة خط انحدار Y على X ، و تعطى بالعلاقة $y = mx + c$ حيث

$$m=\frac{n\sum xy-\sum x\sum y}{n\sum x^2-\left(\sum x\right)^2}=\frac{8\times1322-102\times97}{8\times1398-(102)^2}=0.87$$

$$c=\frac{\sum y}{n}-m\frac{\sum x}{n}=\frac{97}{8}-0.87\times\frac{102}{8}=1.035$$

$$y=0.87x+1.035$$

٢ -معادلة خط انحدار X على Y ، و تعطى بالعلاقة $x=my+c$ حيث

$$m=\frac{n\sum xy-\sum x\sum y}{n\sum y^2-\left(\sum y\right)^2}=\frac{8\times1322-102\times97}{8\times1256-(97)^2}=0.96$$

$$c=\frac{\sum x}{n}-m\frac{\sum y}{n}=\frac{102}{8}-0.96\times\frac{97}{8}=1.11$$

$$x=0.96y+1.11$$

<u>تمارين</u>

١ -إذا كانت لدينا البيانات الآتية

x	0	1	2	3	4	5	6
y	-1	2	2	8	4	14	5

أوجد

١ -معامل الارتباط لبيرسون بين المتغيرين

٢ -معامل الارتباط لسبيرمان بين المتغيرين

٣ -خط انحدار Y على X .

٢ - الجدول الآتي يمثل الدخل X والإنفاق Y بآلاف الريالات

x	56	66	42	44	38	27	39	40
y	31	38	27	22	19	25	20	28

١ ـ أوجد معامل ارتباط بيرسون و سبيرمان

٢ ـ أوجد خط انحدار Y على X

٣ ـ أوجد خط انحدار X على Y

٤ ـ أوجد قيمة الإنفاق إذا كان الدخل 6000 ريال

٣ ـ أوجد معامل الارتباط لتقديرات ثمانية طلاب في مادتي الفيزياء X والكيمياء Y

x	A	B	D	E	C	D	E	B
y	A	C	E	D	C	D	E	B

٤ ـ الجدول الآتي يبين التقديرات التي حصل عليها 480 طالب في اختبارين مختلفين . أوجد معامل التوافق ؟

x/y	A	B	C	$\sum$
C	10	20	100	130
B	30	170	40	240
A	60	30	20	110
$\sum$	100	220	160	480

الباب الخامس

أنظمة العد

مبدأ العد

إذا أمكن إجراء عملية ما بعدد n_1 طريقة مختلفة، ثم تلتها عملية أخرى بعدد n_2 طريقة مختلفة، ... الخ. وهكذا فإن عدد الطرق التي يمكن إجراء العمليات بها بهذا الترتيب هو $n_1, n_2, n_3, \ldots$.

1- مثال

إذا كانت اللوحة المعدنية لرقم سيارة تحتوي على حرفين مختلفين من حروف اللغة العربية يتبعها ثلاثة أرقام بحيث لا يكون الرقم الأول صفر. ما هو عدد اللوحات المعدنية التي يمكن طبعها ؟

الحل

عدد اللوحات المختلفة 28.27.9.10.10=680400

تعريف (رمز المضروب Factorial (n!

إذا كانت n عدد صحيح فإن مضروب n يعرًف كالآتي

$$n! = n(n-1)(n-2)\ldots 3.2.1$$

2 – مثال

$$5! = 5.4.3.2.1 = 120$$

التباديل (Permutation)

يسمى وضع n من الأشياء في ترتيب معين بأنه تبديل لهذه الأشياء – بشرط أن تأخذ جميع هذه الأشياء. ويسمى وضع أي عدد n حيث $r \leq n$ من الأشياء مأخوذة r كل مرة بالتباديل .

3- مثال

لتكن a,b,c,d فأن .

i. bdca, dcba, acdb

هي تباديل من الحروف الأربعة جميعها:

ii. bad, adb, cbd, bca

هي تباديل من الحروف الأربع مأخوذات ثلاث كل مرة.

iii. bd, da, cb, ad

هي تباديل من الحروف الأربع مأخوذات مثنى مثنى.

ملاحظة:-

يرمز لعدد تباديل n من الأشياء مأخوذة r كل مرة بالرمز p(n,r) .

19- مثال:

اوجد عدد الكلمات في اللغة الانجليزية المكونة من ثلاثة أحرف مختلفة من الحروف a,b,c,d,e,f

الحل

عدد الطرق $6 \times 5 \times 4 = 120$

P(6,3)=120

1- نظرية :-

$$\frac{n!}{(n-r)!}$$

البرهان:-

لاشتقاق الصيغة العامة للمقدار p(n,r)

يمكن اختيار العنصر الأول بتبديل n طريقة مختلفة .

يمكن اختيار العنصر الثاني بتبديل(n-1)طريقة مختلفة .

يمكن اختيار العنصر الثالث بتبديل(n-2) طريقة مختلفة .

يمكن اختيار العنصر r بتبديل ((n-(r-1) طريقة مختلفة .

يمكن اختيار العنصر r بتبديل n-r+1 طريقة مختلفة .

$$p(n,r)=n(n-1)...(n-r+1)$$
$$=\frac{n(n-1)(n-2)...(n-r+1)(n-r)!}{(n-r)!}$$
$$=\frac{n!}{(n-r)!}$$

ملاحظة:-

عدد تباديل n من الأشياء مأخوذة جميعها في نفس الوقت هو !n

$$n=r\Rightarrow p(n,n)=n(n-1)...3.2.1=n!$$

<u>4 - مثال</u>

$$p(n,n)=n!=\frac{n!}{(n-n)!}=\frac{n!}{0!}$$
$$\therefore n!=\frac{n!}{0!}$$
$$\therefore 0!=\frac{n!}{n!}=1$$
$$\therefore 0!=1$$

<u>5 - مثال</u>

تباديل العناصر a,b,c مأخوذة جميعها هو 6=3×2×1

والتباديل هي abc, acb, bac, bca, cab, cba التباديل مع التكرار .

<u>2- نظرية</u>

عدد تباديل عنصر والتي من بينها $_{1}$ عنصر متماثلا و n_2 عنصرا متماثلا

و... n_r عنصرا متماثلا هو،

$$\frac{n!}{n_1!n_2!\ldots n_r!}$$

<u>مثال 6</u>

اوجد عدد الكلمات الممكنة من بين تباديل حروف كلمة DADDY

الحل

$$\frac{5!}{3!1!1!}=\frac{120}{6}=20$$

<u>العينات المرتبة</u>

وهي عملية اختيار كرة- مثلا- من وعاء به **n** من الكرات، **r** من المرات وهناك حالتان:

١ - <u>السحب بإرجاع: (with replacement)</u>

عند سحب r كرة من صندوق به n كرة، فإذا سحبت الكرة الأولى ثم أعيدت إلى الصندوق قبل إجراء السحب الثاني. ثم سحبت الكرة الثانية ثم أعيدت قبل السحب الثالث،... ثم سحبت الكرة r ثم أعيدت فإن عدد الطرق التي يتم بها السحب الأول هو n طريقة، عدد الطرق التي يتم بها السحب الثاني هو أيضاً n طريقة عدد الطرق التي يتم بها السحب الثالث هو n طريقة،... ليصبح عدد الطرق التي يتم بها سحب r كرة من صندوق به n كرة مع الإرجاع هي:

$$n\times n\times n\ldots\times n=n^r$$

٢ - السحب بدون إرجاع (without replacement)

عند سحب r كرة من صندوق به n كرة بدون إرجاع فإن حساب عدد الطرق التي يتم بها السحب كالآتي: عدد طرق سحب الكرة الأولى بدون إرجاع هو n طريقة، عدد سحب الكرة الثانية بدون إرجاع هو $n-1$ طريقة عدد الطرق سحب الكرة r بدون إرجاع هو $(n-(r-1))$ طريقة ويصبح عدد طرق سحب r كرة من بين n كرة بدون إرجاع هو:

$$n(n-1)(n-2)\ldots(n-r+1)=p(n,r)$$

<u>مثال -7</u>

وعاء به 15 كرة ما هو عدد طرق سحب كرتين

(i) بإرجاع

(ii) بدون إرجاع.

الحـل

(i) سحب كرتان من بين 15 كرة بإرجاع

$$n^r=(15)^2=225$$

(ii) سحب كرتين من بين 15 كرة بدون إرجاع

$$p(n,r)=p(15,2)=\frac{15!}{(15-2)!}$$

$$=\frac{15!}{13!}=\frac{15\times14\times13!}{13!}=15\times14=210$$

التوافيق: (Combinations)

التوفيق r هو أي اختيار لعدد r شيء من بين n من الأشياء دون اعتبار لعملية الترتيب ويرمز له

$$c\left(n,r\right),or,\binom{n}{r}$$

<u>مثال 23</u>

توافق الحروف a,b,c,d, مأخوذة ثلاثة كل مرة هي 4

$$abc, abd, acd, bcd$$

<u>ملاحظة:</u>

لاحظ التوافيق الآتية كلها متساوية

$$abc, acb, bac, bca, cab, cba$$

إذ لا معنى للترتيب في عملية التوافيق.

<u>مثال 8</u>

سوف نحدد توافيق الحروف الأربعة

$$a, b, c, d$$

مأخوذة 3 كل مرة.

لاحظ أن كل توفيقه مكونة من ثلاثة حروف تناظر **3!=6** تباديل ممكنة للحروف الموجودة في التوفيقة.

$p(n,r)$	$c(n,r)$					
abc	abc	acb	bac	bca	cab	cba
abd	abd	adb	bad	bda	dab	dba
acd	acd	adc	cad	cda	dac	dca
bcd	bcd	bdc	cbd	cdb	dbc	dcb

وبذلك يكون عدد التوافيق مضروبا في **!3** يساوي التباديل :

$$c(4,3).3! = p(4,3)$$

$$c(4,3) = \frac{p(4,3)}{3!} = 4$$

ونستنتج بصورة عامة أن

$$c(n,r) = \frac{p(n,r)}{r!} = \frac{n!}{r!(n-r)!}$$

<u>معادلات ذات الحدين :-</u>

يعطي مفكوك ذات الحدين $(a+b)$

$$(a+b)^n = \binom{n}{0}a^n + \binom{n}{1}a^{n-1}b + \ldots + \binom{n}{r}a^{n-r}b^r + \ldots + \binom{n}{n}b^n$$

حيث معامل الحد العام

$$\binom{n}{r} = \frac{n(n-1)\ldots(n-r+1)}{r(r-1)\ldots 3.2.1}$$

$$= \frac{n(n-1)\ldots(n-r+1)(n-r)!}{r!(n-r)!} = \frac{n!}{r!(n-r)!}$$

$$c(n,r) = \binom{n}{r}$$

9 -مثال

بكم طريقة يمكن اختيار 5 تفاحات من سلة بها 7 تفاحات ؟

الحــل

$$c(n,r) = \binom{n}{r} = \binom{7}{5} = \frac{7!}{5!(7-5)!}$$

$$= \frac{7 \times 6 \times 5!}{5!2!} = 21$$

نظرية ذات الحدين: (Binomial theorem)

لأي عدد صحيح موجب n فإن

$$(x+y)^n = \sum_{r=0}^{n} \binom{n}{r} x^{n-r} y^r$$

حيث

$$\binom{n}{r} = \frac{n!}{r!(n-r)!}$$

تمارين

١ - اوجد عدد الكلمات التي يمكن تكوينها من كلمة Doll.

٢ - اوجد عدد الكلمات التي يمكن تكوينها من كلمة probability .

٣ - بكم طريقة يمكن أن يجلس ثلاثة أشخاص على ثلاثة مقاعد في صف واحد.

٤ - بكم طريقة يمكن اختيار فئة من 4 بنات من فئة مكونة من عشرة بنات.

٥ - 20 شخص تقدموا لوظيفة 15 منهم لائقين طبيا و5 معوقين بكم طريقة يمكن اختيار مجموعة من أربعة أشخاص 3 منهم لائقين وواحد معاق.

٦ - بكم طريقة يمكن ترتيب كرتين حمراء، 3 بيضاء و2 سوداء.

٧ - بكم طريقة يمكن توزيع 7 نزلاء في غرفة لثلاثة أشخاص وغرفتين لشخصين في فندق.

٨ - كم لجنة يمكن تكوينها من كيميائيين وفيزيائي من مجموعة من 4 كيميائي وفيزيائي.

٩ - أوجد عدد طرق اختيار كرتين من صندوق يحتوي على15 كرة بدون ترتيب.

١٠ - اوجد عدد طرق اختيار حرفين من الحروف A,B,C بدون ترتيب.

١١ - موقف مخصص لثمان سيارات في صف واحد بكم طريقة يمكن ترتيب 4 سيارات مرسيدس وسيارة مزدا و 3 تويوتا.

الباب الخامس

مبادئ الإحتمالات

1-**تعريف** التجربة العشوائية

التجربة العشوائية هي التجربة التي تكون جميع نتائجها (مخرجاتها) معلومة مسبقاً ولكن لا يمكن لأحد أن يجزم بحدوث أيٍ من هذه النتائج سلفاً.

1-مثال

عند رمي قطعة نقود فإن جميع نتائجها الممكنة معلومة سلفا وهما ظهور الصورة H أو النقش T لكن لا يستطيع أحد أن يجزم قبل عملية رمي العملة ان الناتج هو الصورة مثلا. عملية رمي قطعة النقود تسمى تجربة عشوائية.

2-**تعريف** فضاء العينة Sample Space

فضاء العينة S هو جميع مخرجات التجربة العشوائية.

2-مثال

فضاء العينة لتجربة رمي قطعة النقود S={H,T}

3-مثال

فضاء العينة لتجربة رمي النرد S={1,2,3,4,5,6}

3-**تعريف** الحادثة Event

الحادثة A هي مجموعة جزئية من فضاء العينة S أي أن $A \subseteq S$

4-مثال

حادثة ظهور الصورة A={H} هي مجموعة جزئية من فضاء العينة S={H,T} أي أن $A \subseteq S$

ملاحظة : فضاء العينة يُسمى الحادثة المطلقة $S \subseteq S$ في حين أن المجموعة الخالية تُسمى الحادثة المستحيلة $\varphi \subseteq S$.

5-مثال

أحسب الحوادث الآتية وعدد عناصر كل جادثة بالنسبة للتجربة العشوائية المتمثلة برمي قطعة النقود مرتين

A = ظهور صورة في الرمية الأولى

A = ظهور نقش في الرمية الأولى

A = ظهور صورة واحدة على الأقل

الحل

$$S=\{HH,HT,TH,TT\},|S|=4$$
$$A=\{HH,HT\},|A|=2$$
$$B=\{TH,TT\},|B|=2$$
$$C=\{HH,HT,TH\},|C|=3$$

ملاحظة :

- الحادثة $A\cup B$ تقع بحدوث A أو B
- الحادثة $A\cap B$ تقع بحدوث A و B معا
- الحادثة A^c - متممة A – تقع إذا لم تحدث A
- $A\cap B=\varphi$ الحادثتان A و B منفصلتان.

4-**تعريف** : التعريف الكلاسيكي للإحتمالات

إذا كان لدينا تجربة عشوائية مخرجاتها متماثلة (متكافئة الفرص كرمي قطعة النقود) وكانت لدينا حادثة $A\subseteq S$ فإن إحتمال حدوث A هو

$$P(A)=\frac{|A|}{|S|}$$

6-مثال

أوجد إحتمال ظهور صورة مرتين عند رمي قطعة نقود مرتين

الحل

$$S=\{HH,HT,TH,TT\},|S|=4$$

$$A=\{HH\},|A|=1$$

$$P(A)=\frac{|A|}{|S|}=\frac{1}{4}$$

7-مثال

أُختير رقم عشوائيا من بين الأرقام من واحد إلى خمسين ما هو أن يكون هذا الرقم 4 أو مضاعفاتها

الحل

$$S=\{1,2,3...,50\},|S|=50$$

$$A=\{4,8,12,...,48\},|A|=12$$

$$P(A)=\frac{|A|}{|S|}=\frac{12}{50}$$

<u>بديهيات الاحتمال</u>

إذا كان S فضاء العينة لكل حادثة A فإن:

$$p_1: 0\le p(A)\le 1$$

$$p_2: p(S)=1$$

$$p_3: p(A\cup B)=p(A)+p(B)$$ A,B منفصلان

<u>1 -نظرية:</u>

لأي حادثتان A,B فإن :

$$(1)P(\phi)=0,(2)P(A^C)=1-P(A)$$
$$(3)A\subset B\Rightarrow P(A)\le P(B),(4)P(A\setminus B)=P(A)-P(A\cap B)$$
$$(5)P(A\cup B)=P(A)+P(B)-P(A\cap B)$$

البرهان:

(١)لأي فئة $A\neq\phi$ فإن $A\cup\phi=A$

وبما أن ϕ , منفصلان و باستخدام P_3 فإن :

$$P(A\cup\phi)=P(A)$$
$$P(A)+P(\phi)=P(A)$$
$$\therefore P(\phi)=0$$

(٢) إذا كان A^C متممة A فإن $S=A\cup A^C$

$$\therefore P(S)=P(A\cup A^C)$$

وبما أن A, A^C منفصلتان فإن :

$$1=P(S)=P(A)+P(A^C)$$
$$\therefore P(A^C)=1-P(A)$$

(٣)بما أن $A\subset B$ فإن B يمكن تجزئته إلى حدثين منفصلين A و A/B

$$\therefore B=A\cup(B\setminus A)$$
$$P(B)=P(A\cup(B\setminus A))$$
$$=P(A)+P(B\setminus A)$$
$$\because P(B\setminus A)\ge 0$$
$$\therefore P(B)\ge P(A)$$

(٤)يمكن تجزئة الحادثة لحدثين منفصلين هما :

$$A = (A \setminus B) \bigcup (AB)$$
$$\therefore P(A) = P(A \setminus B) + P(A \bigcap B)$$
$$\therefore P(A \setminus B) = P(A) - P(A \bigcap B)$$

(٥)يمكن تجزئة الحادثة

$$A \bigcup B = (A \setminus B) \bigcup (B)$$
$$\therefore P(A \bigcup B) = P(A \setminus B) + P(B)$$
$$= P(A) - P(A \bigcap B) + P(B)$$
$$\therefore P(A \bigcup B) = P(A) + P(B) - P(A \bigcap B)$$

8-مثال

أفرض أن A وB حادثتان بحيث أن

$$P(A) = \frac{3}{8}; P(B) = \frac{1}{2}; P(A \bigcap B) = \frac{1}{4}$$

Find

$$(1): P(A \bigcup B); (2): P(A^c); (3): P(B^c)$$
$$(4): P(A^c \bigcup B^c); (5): P(A \bigcap B^c); (6): P(A^c \bigcap B)$$

الحل

$$(1) P(A \cup B) = P(A) + P(B) - P(A \cap B) = \frac{3}{8} + \frac{1}{2} - \frac{1}{4} = \frac{5}{8}.$$

$$(2) P(A^c) = 1 - P(A) = 1 - \frac{3}{8} = \frac{5}{8}.$$

$$(3) P(B^c) = 1 - P(B) = 1 - \frac{1}{2} = \frac{1}{2}.$$

$$(4) P(A^c \cap B^c) = P(A \cup B)^c = 1 - P(A \cup B) = 1 - \frac{5}{8} = \frac{3}{8}.$$

$$(5) P(A \cap B^c) = P(A) - P(A \cap B) = \frac{3}{8} - \frac{1}{4} = \frac{1}{8}.$$

$$(6) P(A^c \cap B) = P(B) - P(A \cap B) = \frac{1}{2} - \frac{1}{4} = \frac{1}{4}.$$

10-مثال

لنفرض أننا اخترنا وحدتين عشوائيا من صندوق به 12 وحدة من بينها 4 وحدات تالفات. ما هو احتمال أن الوحدتين سليمتان، تالفتان أو على الأقل إحداهما تالفة؟

الحل

A=الوحدتان تالفتان ، **B**=الوحدتان سليمتان ، **C**= إحداهما تالفة على الأقل.

$$|A| = \binom{4}{2}; |B| = \binom{8}{2}; |S| = \binom{12}{2}$$

$$P(A) = \frac{|A|}{|S|} = \frac{\binom{4}{2}}{\binom{12}{2}} = \frac{6}{66} = \frac{1}{11}$$

$$P(B) = \frac{|B|}{|S|} = \frac{\binom{8}{2}}{\binom{12}{2}} = \frac{28}{66} = \frac{14}{33}$$

$$\because C = B^c$$

$$\therefore P(C) = P(B^c) = 1 - P(B) = 1 - \frac{14}{33} = \frac{19}{33}.$$

2-تعريف

الحادثتان $A, B \subseteq S$ مستقلتان إذا كان $P(A \cap B) = P(A)P(B)$

11-مثال

عند رمي قطعة نقود 3 مرات وكانت لدينا الحوادث

$$S = \{HHH, HHT, HTH, HTT, TTT, TTH, THT, THH\}; |S| = 8.$$

$$A = \{HHH, HHT\}; |A| = 2, P(A) = \frac{2}{8} = \frac{1}{4}.$$

$$B = \{HHT, HTT, THT, TTT\}; |B| = 4, P(B) = \frac{4}{8} = \frac{1}{2}.$$

$$C = \{HTT, THT, TTH\}; |C| = 3, P(C) = \frac{3}{8}.$$

الحل

$$A \cap B = \{HHT\}$$

$$P(A \cap B) = \frac{1}{8}$$

$$P(A)P(B) = \frac{1}{4} \times \frac{1}{2} = \frac{1}{8}$$

$$\therefore P(A \cap B) = \frac{1}{8} = P(A)P(B)$$

$$B \cap C = \{HTT, THT\}$$

$$P(B \cap C) = \frac{2}{8} = \frac{4}{16}$$

$$P(B)P(C) = \frac{1}{2} \times \frac{3}{8} = \frac{3}{16}$$

$$\therefore P(B \cap C) \neq P(B)P(C)$$

تمارين

1-إذا كان احتمال نجاح محمد 4/1 واحتمال رسوب أحمد 3/1 واحتمال نجاح محمد وأحمد 6/1 فما هو احتمال نجاح أحدهما على الأقل؟

2-إذا كان احتمال نجاح محمد 8/5 واحتمال نجاح محمد وأحمد 8/1 فما هو احتمال نجاح محمد ورسوب أحمد؟

3-اختيرت 3 مصابيح كهربائيةعشوائيا من بين 15 مصباح منها 5 مصابيح تالفة، ما هو احتمال:

أ-جميعها سليمة ، ب-فقط واحد منها تالف ، ج-على الأقل واحد منها تالف

4-نرد صُمم بحيث أن ظهور عدد فردي هو ضعف ظهور عدد زوجي. ما هو احتمال:

أ-ظهور عدد أكبر من 3 في رمية واحدة للنرد

ب-أن الرقم الذي يظهر مربع كامل

ج-الرقم الذي يظهر يكون مربع كامل وأكبر من 3 ؟

5-أختيرت ورقة كتشينة عشوائيا، ما هو احتمال أن تكون ورقة بستوني؟ صورة؟ أو صورة بستوني؟

6-سُحبت ورقتان عشوائيا من كتشينة ما هو احتمال أن كلاهما بستوني؟ إحداهما بستوني ولأخرى قلب؟

7-إذا كانت A و B حادثتان مستقلتين ، برهن أن الحوادث الآتية مستقلة

$$(1) A \quad and \quad B^c$$

$$(2) A^c \quad and \quad B$$

$$(3) A^c \quad and \quad B^c$$

الباب السادس

التوقع الرياضي

المتغيرات العشوائية (Random variable)

المتغير العشوائي المنفصل:

١ -تعريف:

إذا كان **S** فضاء العينة و **X** دالة قيمة حقيقية:

$$X : S \to R$$

بحيث تكون الصورة العكسية لأي فترة تنتمي إلى المدى حادثة في فضاء العينة **S. X** يسمى متغير عشوائي. ويسمى المتغير العشوائي منفصلاً **Discrete** إذا كان مداه منتهي أو مرقم . **Countable** في حين يسمى متصلا ــ CONTINUOUS إذا كان مداه فئة غير منتهية من الأعداد الحقيقية.

٢ -تعريف: داله الكتلة الاحتمالية:

إذا كان **X** متغير عشوائي منفصل فإن الدالة $f(x) = P(X = x)$

-لكل قيم **X** المعرفة ــ تسمى دالة الكتلة الاحتمالية Probability mass function

1- نظرية:

تعمل **f(x)** كدالة كتالة احتمالية للمتغير العشوائي المنفصل فقط إذا كانت **f(x)** تحقق الشرطان **:**

$$1 - f(x) \geq 0$$

$$2 - \sum_{x} f(x) = 1$$

$$\forall x \in X$$

1-مثال:

تحقق ما إذا كانت X=1,2,3,4,5 ، $f(x)=\frac{x+2}{25}$

تعمل كدالة كتلة احتمالية للمتغير العشوائي **X** .

الحـل

$$1-f(1)=\frac{3}{25}>0, f(2)=\frac{4}{25}>0, f(3)=\frac{5}{25}>0,$$

$$f(4)=\frac{6}{25}>0, f(5)=\frac{7}{25}>0,$$

$$2-\sum_x f(x)=f(1)+f(2)+f(3)+f(4)+f(5)$$

$$=\frac{3}{25}+\frac{4}{25}+\frac{5}{25}+\frac{6}{25}+\frac{7}{25}=1$$

ويصبح جدول التوزيع الاحتمالي كالآتي :

X_i	1	2	3	4	5
$f(x_i)=P(x_i)$	3/25	4/25	5/25	6/25	7/25

3-تعريف : دالة التوزيع التراكمي

Cumulative Distribution function

إذا كان **X** متغير عشوائي منفصل فإن الدالة

$$F : R \to R$$
$$F(a) = p(X \leq a)$$
$$= \sum_{t \leq x} f(t)$$

تسمى دالة التوزيع التراكمي .حيث **f(t)** هي دالة الكتلة الاحتمالية لــ **X** عند **t**

2-مثال:

أوجد دالة التوزيع التراكمي للجدول الآتي:

X_i	1	2	3	4
$f(x_i)$	1/4	1/8	1/2	1/8

الحـل

$$F(1) = p(x \leq 1) = p(1) = 1/4$$
$$F(2) = p(x \leq 2) = p(1) + p(2) = \frac{1}{4} + \frac{1}{8} = \frac{3}{8}$$
$$F(3) = p(x \leq 3) = p(1) + p(2) + p(3) = 7/8$$
$$F(4) = p(x \leq 4) = p(1) + p(2) + p(3) + p(4) = 1$$

وعليه يصبح جدول دالة التوزيع التراكمي **F(x)** كالآتي :

X_i	1	2	3	4
$F(x_i)$	1/4	3/8	7/8	1

2-نظرية:

قيمة **F(X)** دالة التوزيع التراكمي – للمتغير العشوائي المنفصل **X** تحقق

الشروط التالية :

$$1)F(-\infty)=0, 2)F(\infty)=1$$
$$3)a<b \Rightarrow f(a)\leq f(b)$$
$$4)f(x_i)=F(X_i)-F(X_{i-1})$$
$$5)p(x>X_i)=1-F(x_i), i=1,2,...,n$$
$$6)p(x\geq x_i)=1-F(x_{i-1})$$

3-مثال:

إذا كانت دالة التوزيع التراكمي للمتغير العشوائي المنفصل كالتالي:-

$$F(x)=\begin{cases}0,...,x<-1\\1/4,...,-1\leq x<1\\1/2,...,1\leq x<3\\3/4,...,3\leq x<5\\1,...,x\geq 5\end{cases}$$

أوجد

$$1)p\left(x \le 3\right), 2)p\left(x = 3\right), 3)p\left(x < 3\right)$$
$$4)p\left(x \ge 1\right), 5)p\left(-0.4 < x < 4\right)$$
$$6)p\left(x = 5\right)$$

الحـل

$$1)p\left(x \ge 3\right) = F\left(3\right) = \frac{3}{4}$$
$$2)p\left(3\right) = F\left(3\right) - F\left(1\right) = 3/4 - 1/2 - 1/4$$
$$3)p\left(X < 3\right) = 1/2$$
$$4)p\left(x \ge 1\right) = 1 - F\left(x < 1\right) = 1 - 1/4 = 3/4$$
$$5)p\left(-0.4 < X < 4\right) = F\left(4\right) - F\left(0.4\right)$$
$$= 3/4 - 1/4 = 1/2$$
$$6)p\left(x = 5\right) = F\left(5\right) - F\left(3\right) = 1 - 3/4 = 1/4$$

4-مثال:

إذا كانت دالة الكتلة الاحتمالية $f\left(x\right) = \frac{x+2}{25}, x = 1,...,5$

برهن أن دالة التوزيع التراكمي تعطى $F\left(x\right) = \frac{x^2 + 5x}{50}$

الحـل

$$F(0)=0, F(1)=\frac{1+5}{50}=\frac{3}{25}, F(2)=\frac{(2)^2+10}{50}=\frac{7}{25},$$
$$F(3)=\frac{9+15}{50}=\frac{12}{25}, F(4)=\frac{16+20}{50}=\frac{18}{25},$$
$$F(5)=1$$

من جهة أخرى فإن

$$F(0)=f(0)=0, F(1)=P(x\le 1)=\frac{1+2}{25}=3/25$$
$$F(2)=P(x\le 1)+P(x=2)=\frac{3}{25}+\frac{4}{25}=7/25$$
$$F(3)=P(x\le 3)=P(x\le 2)+P(x=3)=\frac{7}{25}+\frac{5}{25}=12/25$$
$$F(4)=P(x\le 4)=\frac{18}{25}, F(5)=P(x\le 5)=1$$

المتغير العشوائي المتصل:-

الدالة **f(x)** المعرفة على فئة الأعداد الحقيقية تسمى دالة كثافة احتمالية للمتغير

العشوائي المتصل **X** فقط إذا كان

$$P(a\le x\le b)=\int_a^b f(x)dx : a\le b$$

3-نظرية:

تعمل **f(x)** كدالة كثافة احتمالية للمتغير العشوائي المتصل **X** فقط إذا تحقق

الشرطان:

$$1-f(x)\geq 0,-\infty < x < \infty$$

$$2-\int_{-\infty}^{\infty} f(x)dx = 1$$

5-مثال:

أوجد قيمة **K** التي تجعل الدالة

$$f(x)=\begin{cases} ke^{-3x}, x>0 \\ 0,.....,else \end{cases}$$

تعمل كدالة كثافة احتمالية ثم أوجد P(0.5<x<1)

الحــل

$$\int_{-\infty}^{\infty} f(x)dx = \int_{-\infty}^{\infty} ke^{-3x}dx = k\left.\frac{e^{-3x}}{-3}\right|_{0}^{\infty} = 1$$

$$\therefore k = 3$$

$$p(0.5 < x < 1) = \int_{0.5}^{2} 3e^{-3x} = -e^{-3x}\Big|_{0.5}^{1}$$

$$= -e^{-3} + e^{-15} = 0.173$$

4-تعريف: دالة التوزيع التراكمي:

إذا كان **X** متغير عشوائي متصل فإن دالة التوزيع التراكمي كالآتي

$$F(x)=P(X \le x)=\int_{-\infty}^{t} f(t)dt, -\infty < x < \infty$$

$$F(-\infty)=0, F(\infty)=1$$

4-نظرية:

إذا كانت **f(x)** دالة الكثافة الاحتمالية و **F(x)** دالة التوزيع التراكمي للمتغير

العشوائي المتصل **x** فإن

$$f(x)=P(a<x\le b)=F(b)-F(a)$$

$$f(x)=\frac{dF(x)}{dx}$$ **حيث**

إذا وجدت المشتقة

6-مثال:

أوجد دالة التوزيع التراكمي المناظرة لدالة الكثافة الاحتمالية $f(x)=3\,e^{-3x}$

x,>0

ثم أوجد $P(0.5 \le x \le 1)$

الحـل

$$F(x)=\int_{-\infty}^{x} f(t)dt=\int_{0}^{x} 3e^{-3t}dt$$

$$=-e^{-3t}\Big|_{0}^{x}=1-e^{-3x}$$

$$\therefore F(x)=\begin{cases}0, x\leq 0\\ 1-e^{-3x}, x>0\end{cases}$$

$$\therefore p(0.5\leq x\leq 1)=F(1)-F(0.5)$$

$$=(1-e^{-3})-(1-e^{-1.5})=0.173$$

التوقع الرياضي (Mathematical Expectation):

5-تعريف التوقع الرياضي:

(I) التوقع الرياضي للمتغير العشوائي المنفصل x الذي دالة كتلته الاحتمالية f(x) هو

$$\mu=E(x)=\sum_{x} xf(x)$$

(II) التوقع الرياضي للمتغير العشوائي المتصل x الذي دالة كتلته الاحتمالية f(x) هو

$$\mu=E(x)=\int_{-\infty}^{\infty} xf(x)dx$$

حيث E(x) هو الوسط المرجح للقيم الممكنة للمتغير العشوائي وفيزيائيا فإن التوقع الرياضي يساوي مركز الثقل.

7-مثال:

إذا كانت دالة الكتلة الاحتمالية للمتغير العشوائي المنفصل X معطاة وفق الجدول التالي :

x_i	1	2	3	4	5	6
$f(x_i)$	1/36	3/36	5/36	7/36	9/36	11/36

أوجد التوقع الرياضي لــــلمتغير X

الحــل

$$E(x) = \sum x_i f(x_i)$$
$$= 1.\frac{1}{36} + 2.\frac{3}{36} + 3.\frac{5}{36} + 4.\frac{7}{36} + 5.\frac{9}{36} + 6.\frac{11}{36}$$
$$= \frac{161}{36} = 4.47$$

8-مثال:

أوجد التوقع الرياضي للمتغير العشوائي المتصل x الذي دالة كثافته

$$f(x) = \begin{cases} \frac{4}{\pi(1+x^2)}, 0 < x < 1 \\ 0, ..., else \end{cases}$$

الحــل

$$E(x)=\int_{-\infty}^{\infty} xf(x)=\int_{0}^{1} x.\frac{4}{\pi\left(1+x^{2}\right)}dx$$

$$=\frac{4}{\pi}\int_{0}^{1}\frac{x}{1+x^{2}}dx=\frac{2}{\pi}In\left(1+x^{2}\right)\Big|_{0}^{1}$$

$$=\frac{2}{\pi}In(1+1)=\frac{2}{\pi}In2=\frac{In4}{\pi}=0.441$$

التباين الرياضي

6-تعريف التباين:

يعرف تباين المجتمع الذي حجمه n بأنه

$$V(x)=\frac{1}{n}\sum_{i=1}^{n}\left(x_{i}-\bar{x}\right)^{2}=E\left[\left(x-E(x)\right)^{2}\right]$$

حيث n حجم المجتمع ومتوسط المجتمع هو $\bar{x}=\frac{1}{n}\sum_{i=1}^{n}x_{i}$

5-نظرية:

$$V(x)=E\left[\left(x-E(x)\right)^{2}\right]=E\left(x^{2}\right)-\mu^{2}$$

البرهان

$$V(x)=E\left[\left(x-E(x)\right)^{2}\right]=E\left[\left(x-(\mu)\right)^{2}\right]^{2}$$

$$=E\left(x^{2}-2\mu x+\mu^{2}\right)=E\left(x^{2}\right)-2\mu E(x)+E\left(\mu^{2}\right)$$

$$=E\left(x^{2}\right)-2\mu^{2}+\mu^{2}=E\left(x^{2}\right)-\mu^{2}$$

9-مثال:

أوجد التباين الرياضي للمتغير العشوائي المنفصل x في المثال السابق

الحـل

لدينا من المثال السابق $E(x) = \mu = 4.47$

$$E\left(x^2\right) = \sum x_i^2 f\left(x_i\right)$$

$$= 1^2.\frac{1}{36} + 2^2.\frac{3}{36} + 3^2.\frac{5}{36} +$$

$$4^2.\frac{7}{36} + 5^2.\frac{9}{36} + 6^2.\frac{11}{36}$$

$$= \frac{791}{36} = 21.97$$

$$\therefore V\left(x\right) = E\left(x^2\right) - \mu^2 = 21.97 - 19.98 = 1.99$$

10-مثال:

أوجد التباين للمتغير العشوائي المتصل x في المثال السابق

الحـل

لدينا من المثال السابق :

$$\mu = E(x) = 0.441$$

$$E(x^2) = \int_0^1 x^2 \frac{4}{\pi(1+x^2)} dx = \frac{4}{\pi}\int_0^1 \left(1 - \frac{1}{1+x^2}\right) dx$$

$$= \frac{4}{\pi}\left(x - \tan^{-1} x\right)\Big|_0^1 = \frac{4}{\pi} - 1 = 0.273$$

$$\therefore \sigma^2 = v(x) = E(x) = E(x^2) - \mu^2 = (0.273) - (0.441)^2$$

$$= 0.0785$$

الانحراف المعياري هو الجذر التربيعي للتباين

$$\sigma = \sqrt{v(x)}$$

ملاحظة:

فيزيائياً هو عزم القصور الذاتي للنظام.

6-نظرية:

$$v(ax + b) = a^2 v(x)$$

البرهان

$$v(ax+b)=E\left[(ax+b)-E(ax+b)^2\right]$$

$$=E\left[(ax+b)-(aE(x)+E(b))^2\right]$$

$$=E\left[(ax-aE(x))^2\right],E(b)=b$$

$$E\left[a^2(x-E(x))^2\right]=a^2E\left[(x-E(x))^2\right]$$

$$=a^2v(x)$$

<u>الدوال الاحتمالية المشتركة</u>

<u>7-نظرية:</u>

الدالة الثنائية (X,Y) تعمل كدالة كتلة احتمالية مشتركة للمتغيرين المنفصلين X,Y فقط إذا كان

1) $f(x,y)\geq 0$ 2) $\sum_x\sum_y f(x,y)=1$

(ii) الدالة الثنائية $f(x,y)$ تعمل كدالة كثافة احتمالية مشتركة للمتغيرين العشوائيين المتصلين y,x إذا فقط تحقق الشرطان

$$\int_{-\infty}^{\infty}\int_{1\infty}^{\infty} f(x,y)dy \quad (1$$

$$f(x,y)\geq 0 \quad \begin{matrix}\infty<x<\infty\\ -\infty<y<\infty\end{matrix}(2$$

11-مثال:

أوجد قيمة K التي تحقق أن الدالة

تعمل كدالة كتلة احتمالية $f(x,y)=kxy \quad \begin{matrix} x=1,2,3 \\ y=1,2,3 \end{matrix}$

الحــل

$$\sum_x \sum_y f(x,y)=k\sum_x \sum_y xy$$

$$=f(1.1)+f(1.2)+f(1.3)+f(2.1)+(2.2)+f(2.3)$$

$$+f(3.1)+f(3.2)+f(3.3)$$

$$=k+2k+3k+2k+4k+6k+3k+6k+9k=36k=1$$

$$\therefore k= 1/36$$

12-مثال:

إذا كانت دالة الكثافة الاحتمالية المشتركة هي:-

$$f(x,y)=\begin{cases} \frac{3}{5}\times(x,y) & 0<x<1<0<y<2 \\ 0 & \text{عدا ذلك} \end{cases}$$

أوجد $p(0<x<1/2, 1<y<2)$

الحــل

$$p\left(0<x<\tfrac{1}{2},1<y<2\right)=\int_1^2\int_0^{\frac{1}{2}}\frac{3}{5}x\left(y+x\right)dxdy$$

$$=\int_1^2\left(\frac{3x^2}{10}y+\frac{3x^2}{15}\right)_0^{\frac{1}{2}}dy=\int_1^2\left(\frac{3y}{40}+\frac{1}{40}\right)dy$$

$$=\left(\frac{3y^2}{80}+\frac{y}{40}\right)_1^2=\frac{11}{80}$$

13-مثال

إذا كانت دالة الكتلة الاحتمالية المشتركة معرفة وفقا للجداول التالية:

أوجد

$Y \backslash X$	0	1	2
0	1/6	1/3	1/12
1	2/9	1/6	
2	1/36		

$(1)p\left(x\le 1,y\le 1\right)$

$(2)F\left(-2.1\right)$

$(3)F\left(3.7,4.5\right)$

الحـل

$$(1)F\left(-1.1\right)=p\left(x\le 1,y\le 1\right)f\left(0.0\right)+f\left(0.1\right)$$

$$+F\left(1.0\right)+f\left(1.1\right)=\frac{1}{6}+\frac{2}{9}+\frac{1}{3}+\frac{1}{6}=\tfrac{8}{9}$$

$$(2)F\left(-2.1\right)=p\left(x\le -2,y\le 1\right)=0$$

$$(3)F\left(3.7,4.5\right)=P\left(X\le 3.7,Y\le 4.5\right)=1$$

14-مثال:

إذا كانت دالة لتوزيع التراكمي المشتركة للمتغيرين المتصلين X,Y هي

$$F(x,y)=\begin{cases}\left(1+e^{-x}\right)^{y} & x>0, y>0\\ 0 & else\end{cases}$$

أوجد $p(1<x<3, 1<y<2)$

الحـل

$$f(x,y)=\begin{cases}\dfrac{\partial^2 F(x,y)}{\partial x\,\partial y}=e^{-(x+y)} & x>0, y>0\\ 0 & else\end{cases}$$

$$\therefore p(1<x<3, 1<y<2)=\int_1^2\int_1^3 e^{-(x+y)}dxdy$$

$$=\left(e^{-1}-e^{-3}\right)\left(e^{-1}-e^{-2}\right)=0.074$$

14-مثال:

أوجد قيمة C التي تجعل الدوال كتلة احتمالية:

$$(1)f(x)=cx^2 \quad (2)f(x)=c2^x \quad (3)f(x)=cx^2$$

$$x=1,2,3....,k \quad x=1,2,.....,k \quad x=1,2,3,.....$$

$$(4) f\ (1-c)c^2 \qquad x = 0,1,2,.....$$

الحـل

$$(1) \sum f\ (x) = \sum_{x=1}^{k} cx^2 = c\sum_{x=1}^{k} x^2 = c.\frac{k}{6}(k+1)(2k+1) = 1$$

$$\therefore C = \frac{6}{k(k+1)(2k+1)}$$

$$(2) \sum f\ (x) = \sum_{x=1}^{k} C\,2^x = C\left(2+2^2+3^2+...+2^k\right) = 1$$

$$= C\,2\frac{\left(1-2^{k-1}\right)}{1-2} = 2C\left(2^{k-1}-1\right).\therefore \frac{1}{2^k-2}$$

$$(3) \sum_x f\ (x) = C\sum_{x=1}^{\infty}\left(\frac{1}{4}\right)^x = \frac{c}{4}.\frac{1}{1-\frac{1}{4}} = 1 \therefore C = 3$$

$$(4) \sum_{x=0}^{\infty} f\ (x) = (1-C)\sum_{x=0}^{\infty} C^x = (1-C)\left(C^0+C+C^2+...\right)$$

$$= (1-C)\frac{1}{1-C} = 1 \therefore 0 < c < 1$$

<u>15-مثال:</u>

أوجد قيمة C التي تجعل الدوال التالية دوال كثافة احتمالية

$$(1) f\ (x) = \begin{cases} c/\sqrt{x} & 0 \le x \le 4 \\ 0 & else \end{cases}$$

$$(2) f\ (z) = \begin{cases} Ce^{-z^2} & z > 0 \\ 0 & else \end{cases}$$

الحـل

$$(1)\int_{-\infty}^{\infty} f(x)dx = c\int_0^4 \frac{dx}{\sqrt{x}} 2Cx^{½}\Big|_0^4 = 4C = 1$$

$$\therefore C = \frac{1}{4}$$

$$(2)\int_{-\infty}^{\infty} f(z)dz = C\int_0^{\infty} ze^{-z^2}dz = \frac{C}{2}0\int_0^{\infty} e^{-z^2}dz^2$$

$$= -\frac{C}{2}e^{-z^2}\Big|_{-\infty}^{\infty} = -\frac{C}{2}[0-1] = \frac{C}{2} = 1$$

$$\therefore C = 2$$

16-مثال:

برهن أن $p(a \le x \le b) = p(a < x < b)$

حيث X متصل

الحـل

$$p(a \le x \le b)\int_a^a f(x)dx + \int_x^b f(x)dx + \int_x^b f(x)dx + \int_b^b f(x)dx$$

$$= 0 + \int_a^x fxdx + \int_x^b f(x)dx + 0 = p(a < x < b)$$

17-مثال:

أوجد التوقع $\mu = E(x)$ والتباين σ^2 والانحراف المعياري للتوزيع التالي:-

x_i	2	3	11
$f(x_i)$	⅓	½	⅙

الحـل

$$(1)\mu = E(x) = \sum xif(x_i) = 2\times\frac{1}{3} + 3\times\frac{1}{2} + 11\times\frac{1}{6} = 4$$

$$(2)E(x^2) = \sum x_1^2 f(x_i) = 2^2\times\frac{1}{3} + 3^2\times\frac{1}{2} + 11^2\times\frac{1}{6} = 26$$

$$\therefore \sigma^2 = E(x^2) - \mu^2 = 26 - 16 = 10$$

$$\therefore \sigma = \sqrt{\sigma^2} = \sqrt{10} = 3.2$$

<u>أمثلة للإحتمالات في حياتنا اليومية</u>

<u>التوقع الرياضي والمخاطرة</u>

في حالة أن المتغير العشوائي x هو الربح العشوائي في لُعبة قابلة للتكرار تحت نفس الشروط ، فإن التوقع الرياضي للمتغير العشوائي x هو معلومة تفيد عملية القياس واتخاذ القرار مبنية على أساس قانون الأعداد الكبيرة.

<u>مثال-18</u>

في لعبة تسمى (فوق وتحت) يتم رمي نردان ويمكنك أن تراهن على أن مجموعهما " تحت 7 – تكسب ريال مقابل ريال إذا فزت- " ، " فوق 7-تكسب ريال مقابل ريال إذا فزت- " أو "يساوي سبعة – تكسب 4 ريالات مقابل ريال إذا فزت-" . أي لاعب يستطيع أن يراهن على حالة أو اكثر من الحالات السابقة. إذا كانت استراتيجيتك أن تراهن بريال على " تحت 7 " وريال آخر على " فوق 7 " كل جولة . ما هو متوسط الربح أو الخسارة في الجولة إذا لعبت عدد كبير من الجولات ؟

الحل

ليكن المتغير العشوائي x يمثل عدد النقود التي سوف تستعيدها في أي جولة .

القيم الممكنة هي 0 و 1 و 5 . هنالك 6 حالات يكون فيها المجموع 7 و 15

حالة يكون فيها المجموع فوق 7 و 15 حالة يكون فيها المجموع تحت 7 .وهذا يعني أن

$$P(x=0)=\frac{15}{36}, P(x=2)=\frac{15}{36}, P(x=5)=\frac{6}{36}$$

$$E(x)=0\times\frac{15}{36}+2\times\frac{15}{36}+5\times\frac{6}{36}=1\frac{2}{3}$$

هذا يعني عندما تراهن بريالين في كل جولة فإن معدل ما تخسره كل جولة هو المبلغ

$$2-1\frac{2}{3}=\frac{1}{3}$$

19-مثال مسألة أفضل اختيار

لنفرض أن صديقك تحداك على الرهان الآتي : عشرون بطاقة مبعثرة عشوائياً فوق منضدة كل بطاقة تحمل رقم مختلف مخفي بداخلها . سيفتح صديقك بطاقة تلو الأخرى وعليه أن يقرر فيما سيتوقف إذا حدس أن هذه البطاقة تحمل أكبر رقم أم يستمر فيما عدا ذلك . سيدفع لك ريال إذا لم تحمل تلك البطاقة أكبر رقم وفي المقابل ستدفع له 5 ريالات إذا صح حدسه ، هل ستقبل الرهان أم لا ؟

الحل

لنفرض أن الإستراتيجية التي سيتّبعها صديقك على النحو الآتي : سيمر صديقك على العشرة بطائق الأولى فقط مرور الكرام مع تدوين أكبر رقم من بينها كملاحظة في عقله . أثناء تفحصه ما تبقى من البطائق العشرة الأخيرة سيتوقف عند أول بطاقة تحمل رقم أكبر من ذلك الذي يحتفظ به في ذاكرته . مؤكد أن هذه البطاقة ستظهر فقط إذا لم يكن أكبر رقم من بين أرقام البطائق العشرة الأولى .

ليكن p هو هو احتمال – مجهول – أن يفوز صديقك بالرهان . التوقع الرياضي لصافي مكسبك سيكون

$$(1-p)\times 1-p\times 5=1-6p$$

واضح أن سياق الرهان ليس في مصلحتك إذا كانت $p>\frac{1}{6}$. لنتخيل أن البطاقة التي تحمل أكبر رقم معلّمة بحبرٍ سري بالرقم 1 والتي تليها بالرقم 2 وهلم جرا . أصبح لدينا 20 بطاقة تحمل الأرقام الخفية $1,2,3,\ldots,20$ وعدد التباديل الممكنة لها $20!$ والتي تمثل فضاء العينة . ليكن A حادثة أن الرقم الخفي 2 يقع ضمن العشرة بطائق الأولى ، في الوقت الذي تقع فيه البطاقة التي تحمل الرقم -الخفي 1-الأكبر ضمن العشرة بطائق الأخيرة . لإيجاد الإحتمال $p(A)$ فإن

$$p(A)=\frac{10\times 10\times 18!}{20!}=\frac{100}{20\times 19}=0.263>25\%.$$

20-مثال مسألة عيد الميلاد

مسألة عيد الميلاد مسألة شهيرة في حقل الإحتمالات . في مباراة كرة القدم يلعب 22 لاعب إضافة إلى حكم المباراة ، ما هو احتمال أن 2 منهم يحتفلان بعيد ميلادهما في نفس اليوم ؟

الحل

بشكل عام نعيد طرح السؤال كالآتي : ما هو احتمال أن شخصين من مجموعة مختارة عشوائياً من **n** شخص يتفقان في نفس يوم عيد ميلادهما ؟ . سنحسب أولاً الإحتمال المتمم ألا وهو احتمال أن لا يوجد 2 لهما نفس يوم عيد الميلاد . لنفرض أن الأشخاص مرقمون : **1,2,...,20** ومن ثم يوجد 365^n حالة ممكنة لأعياد ميلاد لمجموعة مرتبة من n شخص . عدد الحالات التي لا يتفق فيها اثنان في يوم عيد الميلاد تساوي

$$365\times 364\times\ldots\times(365-n+1)$$

احتمال أن لا يوجد 2 لهما نفس يوم عيد الميلاد هو

$$\frac{365\times364\times...\times\left(365-n+1\right)}{365^{n}}$$

ومن ثم فإن احتمال أن شخصين من مجموعة مختارة عشوائياً من **n** شخص يتفقان في نفس يوم عيد ميلادهما هو

$$P_n=1-\frac{365\times364\times...\times\left(365-n+1\right)}{365^{n}}$$

ولبعض قيم **n** نحصل على الجدول الآتي

n	15	20	23	25	30	40	50	75
p_n	0.25	0.41	0.51	0.57	0.71	0.89	.97	0.99

تمرين : برهن لقيم **n** الكبيرة فإن

$$p_n\approx1-e^{-\frac{1}{2}\frac{n(n-1)}{365}}$$

21-مثال البيانات المزورة

معظم الناس لديهم مفاهيم مسبقة عن العشوائية التي تختلف في كثير من الأحيان إلى حد كبير من العشوائية الحقيقية . البيانات العشوائية الحقيقية في حقيقة الأمر تتميز بخواص عادةً لا تتفق مع نمط التفكير البديهي . هذه الخواص يمكن اللجوء إليها لإختبار إذا تبادر الشك أن العبث قد تم بمجموعة البيانات العشوائية الحقيقية. هب أن شخصين طُلب من كلٍ منهما على حِدة أن يرمي قطعة نقود 100 رمية متتالية ويرصد النتائج حيث يرمز 1 للرأس ويرمز 0 للذيل .كانت النتائج كالآتي

1 1 0 0 1 0 0 1 0 1 1 0 0 1 0 0 0 1

1 0 1 0 0 1 1 0 1 0 0 1 0 1 0 1 1 0

0 0 1 1 0 1 1 1 0 1 0 0 1 0 0 1 1 0

0 1 1 0 1 0 0 1 1 0 1 0 1 1 0 0 1 1

0 1 0 1 0 1 0 0 0 1 0 1 0 1 0 1 0 1

1 0 0 1 0 0 1 0 1 1 0 0 1 0 0 1 1 0

and

1 1 1 0 0 0 1 1 1 0 1 0 1 1 1 1 1 1

0 0 0 1 1 0 0 1 1 0 1 0 1 0 0 0 1 1

0 0 1 1 1 0 1 0 0 0 0 1 0 1 1 1 0 1

0 1 1 1 0 1 1 0 0 1 1 1 1 1 1 0 1 1

0 1 1 1 0 0 0 0 0 0 0 0 1 1 0 1 1 1

1 1 1 0 1 1 1 1 0 1 0 1 1 0 1 1 0 1

يبدو أن أحدهما قد غش وألف النتائج من بنات أفكاره دون أن يجري تجربة رمي قطعة النقود 120 رمية فعلية ، ياتُرى من هو ؟ . هنالك احتمال كبير جداً خلال عملية الرمي أن يتكرر الرأس (أو الذيل) على التوالي على الأقل 5 مرات أو يزيد وهذا الإحتمال يساوي 0.9865. من هنا يتضح أن الجدول الأول ملفق . لكي نحسب الأحتمال آنف الذكر ، سنطرح السؤال الآتي : ما هو احتمال حدوث متوالية من **r** رأس في **n** رمية قطعة نقود ؟ للإجابة على هذا السؤال سنفترض أن عملية رمي قطعة النقود في الوضع (i,k) عندما لايزال هنالك k رمية متبقية و متوالية من i رأس قد تحققت لكن دون أن تكتمل r . لنعرف $u_k(i)$ بأنه يساوي احتمال حدوث متوالية من **r** رأس في **n** رمية قطعة نقود إذا كان الوضع الحالي لعملية رمي قطعة النقود في الوضع

$$(i,k), k=0,1,...,n.and, i=0,1,...,r$$

ويجري البحث عن الإحتمال $u_n(0)$. لإيجاد المعادلة التكرارية للإحتمال $u_k(i)$ سنشترط أن الحالة التي تلي الوضع (i,k) ، سيكون احتمال حدوث الرأس في

الرمية التالية هو $1/2$. إذا حدث هذا فإن الوضع التالي لعملية الرمي يصبح $(i+1, k-1)$ ؛ مالم فإى الوضع التالي هو $(0, k-1)$. من قانون الإحتمال المشروط نحصل على الصيغة التكرارية الآتية

$$u_k(i) = \frac{1}{2}u_{k-1}(i+1) + \frac{1}{2}u_{k-1}(0)$$
$$k = 1, 2, \ldots\ldots, n$$
$$i = 0, 1, \ldots, r-1$$

الباب السابع

تاريخ الاحتمالات

نظرية الاحتمالات هي لا شئ سوى صياغة للحس المشترك common sense في لغة رياضية .

أصل نظرية الاحتمالات :-

يعتبر ميلاد نظرية الاحتمالات مدين لعمليات المراهنة و لعب الميسر . لقد عرفت البشرية لعب الميسر منذ أقدم عصورها و في مختلف المناطق . كانت الكهانة تمارس ضرباً من الرهان في الطقوس الدينية القديمة . مثلاً إجراء القرعة لاختيار الشخص – تعيس الحظ – الذي سيقدم قرباناً للآلهة الوثنية . الكاهن الوسيط بين الآلهة و البشر و الذي يعبر عن رغبات الآلهة يرمي بالنرد على أرضية المعبد و يتأول الناتج كجواب عن رضا الآلهة من عدمه . لم يكن هناك ما هو عشوائي ، لا يوجد حظ لأن تأثير الآلهة مطبق على كل شئ و إرادتها تحكم مجريات الأمور صغيرها و كبيرها ، إنها الجبرية القاهرة و القدر المحتوم الذي لا يملك الإنسان عنه فكاكاً . حتى النتائج التي تمخضت عن رمي النرد تنسجم مع إرادة الآلهة و تخضع لها . أقدم نرد أثري أكتشف في شمال العراق و يرجع تاريخه إلى 5000 سنة قبل الميلاد . ظهر النرد بوجوهه الستة المألوف لدينا قبل ميلاد المسيح . إبان حصار طروادة (1200 قبل الميلاد) الذي استمر عشر سنوات أبدع الجنود ، الذي دب في قلوبهم السأم ، خلاله العديد من ألعاب الميسر . في عهد الإمبراطورية الرومانية كانت المقامرة بواسطة النرد إحدى وسائل الترفيه و التسلية . الإمبراطور الروماني Claudius (10 ق.م – 54 م) كرّس معظم وقته للمراهنة بالنرد و ألف كتاب بعنوان (كيف تفوز باستخدام النرد) .

على الرغم من انتشار الميسر بين الإغريق و الرومان و استحسانه إلا أنه كان محرم قطعياً عند اليهود و عقوبته الإعدام ، و ذلك لأن المقامر يربح شئ مقابل لا

شئ . يرجع أصل لعب الورق (الكوتشينة) إلى المصريين و الصينيين و الهنود و لم تعرفه أوروبا إلا في مطلع القرن الرابع عشر . في العصور الوسطى شنت الكنيسة المسيحية حملة شعواء ضد اللعب بالنرد و الورق (الكوتشينة) و ذلك لما يخالطهما من سب و شتم و شرب خمر و فسوق و مجون . حرّم لويس التاسع (1214 – 1270) ملك فرنسا اللعب بالنرد و منع صناعته في أرجاء مملكته (1)[1].

الميسر ، أياً كانت أدواته ،بالطبع محرّم في الإسلام بنص القرآن (يا أيها الذين آمنوا إنما الخمر و الميسر و الأنصاب و الأزلام رجسٌ من عمل الشيطان فاجتنبوه لعلكم تفلحون . إنما يريد الشيطان أن يوقع بينكم العداوة و البغضاء في الخمر و الميسر و يصدكم عن ذكر الله و عن الصلاة فهل أنتم منتهون)) المائدة الآيات 90 - 91 إلا أن مجرد اللعب بالنرد أو الورق بدون مقامرة ليس هناك ما يجزم أو يفيد بتحريمه . سبعة آلاف سنة من ممارسة الميسر و اللعب بالنرد و الورق مهدت لبعض إرهاصات نظرية الاحتمالات و ظهور المبادئ الأولية لها . لم يتبلور ارتباط واضح بين المراهنة و الرياضيات إلا في وقت متأخر . ربما لم يكن النرد المستخدم قديماً يتمتع بقدر كافي من الاتزان و الانتظام و تكافؤ للفرص يوحي أو يقود إلى بعض قوانين الحظ و الصدفة و الاحتمال ، إضافة إلى غياب الرموز الرياضية الملائمة أعاق الوصول إلى تلك القوانين . أو قد يكون المبرر المعقول هو أن مفهوم العشوائية و الحظ و الصدفة دخيل على نمط التفكير السائد و المبني على الحقيقة المطلقة حسب تعاليم الكنيسة بل و حتى المعتقدات الكهنوتية العتيقة الجبرية و التي تعتقد و تسلم بإدارة الآلهة المباشر لشئون الأرض . نتيجة لكل هذا و ذاك تأخر كثيرا منهج أو مدرسة تأخذ في حسبانها الحظ و الصدفة . الفكرة السائدة التي تفيد بأن نظرية الاحتمالات بدأت كفرع للرياضيات لأول مرة نتيجة الرسائل المتبادلة بين عالمي الرياضيات باسكال و فيرمات عام 1654 .

إلا أن هذا غير صحيح على وجه الدقة ربما لسنوات عدة خلت قبل أن يفلح كل من لابلاس و فيرمات في تعريف (القيمة الصادقة للحظ) ، عالج بعض الرياضيين العديد من المسائل ذات الطبيعة الاحتمالية . على الرغم من ذلك من الإنصاف الاعتراف بأن الفضل يرجع لجهود باسكال و فيرمات في أعطاء دفعة حيوية ، عبر سلسلة من الأفكار ، مهدت و بلورت لنظرية الاحتمالات على الوجه الذي نألفه اليوم .

كاردان عالم الرياضيات الإيطالي – الذي توصل إلى الحل العام لمعادلة الدرجة الثالثة - كتب عام 1550 ، فيما يعتبر أول بحث يربط بين ألعاب الحظ و مبادئ نظرية الاحتمالات و لم ينشر إلا بعد وفاته عام 1663 وضع أول تعريف للاحتمال ، و إن كان غير دقيق :- (احتمال تحقق ناتج معين هو مجموع كل الطرق الممكنة لحصول ذلك الناتج مقسوم على مجمل الطرق الممكنة) . مثلاً في عملية رمي قطعة نقود فإن مجمل الطرق هما اثنان ، ظهور الصورة أو ظهور النقش ، و يصبح احتمال ظهور الصورة هو 2/1 و هو نفسه مساوٍ للاحتمال الآخر لظهور النقش ، لأن ظهور كل واحدٍ منهما يتمتع بفرصة واحدة فقط .

مثال آخر عند رمي النرد تتكافأ فرص ظهور أي وجه من وجوهه الستة {1, 2, 3, 4, 5, 6} في حين إن مجمل الطرق هو 6 و يصبح مثلاً احتمال ظهور الوجه الذي يحمل الرقم 4 هو 6/1 و احتمال أن الوجه الذي يظهر يحمل رقم زوجي هو 6/3 لأن هناك ثلاثة أوجه تحمل رقم زوجي و هي الاثنان و الأربعة و الستة و بالتالي مجموع الطرق الممكنة 3 في حين أن مجمل كل الطرق هي 6 . يعتبر كاردان الأب الشرعي لنظرية الاحتمالات الحديثة . (1)2

بدأت الاحتمالات كعلم تجريبي ثم تطورت لاحقاً في اتجاه الرياضيات . تفرعت الاحتمالات إلى فرعين :

أولاً : كحلول للمسائل المتعلقة بالحظ في ألعاب الرهان .

ثانياً : معالجة البيانات الإحصائية الخاصة بالتأمين و جداول المواليد و الوفيات .

باسكال (Blaise Pascal 1623-1662) :-

يعتبر باسكال من الآباء المؤسسين لنظرية الاحتمالات . لقد كان باسكال بارع و موهوب في العديد من المجالات ، فهو رياضي مبدع و فيلسوف لاهوتي و أيضاً فيزيائي تجريبي . أحرز باسكال تقدماً كبيراً فيما يتعلق بمعاملات مفكوك ذات الحدين كان من شأنه أن مهد لإرساء قواعد نظرية الاحتمالات . لقد كان العام 1654 علامة فارهة في تاريخ نظرية الاحتمالات ، بل أصبح العام الذي يؤرخ فيه لميلاد النظرية . النبيل الفرنسي دي مير ، الذي كان مولعاً بالمراهنات ، نتيجةً لبعض المسائل المزعجة التي واجهته في عمليات الرهان أرسل بعض الأسئلة مستفسراً من باسكال ، مما أثار فضول باسكال . باسكال بدوره أشرك معه الرياضي الضليع و اللامع في عصره فيرمات ، و من خلال الرسائل المتبادلة بينهما أسس الرجلان القواعد التي شيِّد عليها صرح نظرية الاحتمالات . كان دي مير رجل متقد الذكاء وعلى قدر من المعرفة الرياضية ، و إن لم يكن ضليعاً فيها ، مكنته من إجراء بعض الحسابات الاحتمالية في عمليات الرهان . كتب دي مير إلى باسكال قائلاً : (لقد اكتشفت في الرياضيات أشياء جديدة لم تتطرق لها الرياضيات العتيقة) و يضيف دي مير موضحاً المشكلة التي واجهته : (إذا كان لدينا نرد مكتمل له ستة وجوه مرقمة من واحد إلى ستة و النرد متزن بحيث أن أي وجه من الوجوه يتمتع بنفس الحظ في الظهور ، و هذا يعني أن احتمال ظهور الوجه ستة عند رمي النرد مرة واحدة هو 6/1 و هذا يكافئ أن نقول احتمال عدم ظهور الوجه ستة عند رمي النرد مرة واحدة هو 6/5 . إذا رمينا النرد مرتان فإن هناك 6×6 حالة ممكنة منها 5×5 حالة ممكنة من عدم ظهور الوجه ستة و يصبح احتمال عدم ظهور الوجه ستة في كلا الرميتان هو $(6/5)^2$. و يصبح عدم ظهور الوجه ستة في عدد نون من الرميات هو $(6/5)^{ن}$ ، و عليه فإن الوضع

المعاكس و هو ظهور الوجه ستة على الأقل مرة واحدة يحدث باحتمال 1 – $(6/5)^{ن}$ > 1/2 .

على سبيل المثال لو أخذت ن = 4 فإن طرف المتراجحة الأيمن هو 671/1296 و هو فعلاً أكبر من النصف ، لذا وجد دي مير من المجدي له و المربح أن يراهن على ظهور الوجه ستة مرة واحدة على الأقل في أربع رميات للنرد .

جيمس برنوللي James Bernoulli 1654-1705 :-

ولد برنوللي في نفس العام الذي شهد مولد نظرية الاحتمالات نتيجة الرسائل المتبادلة بين باسكال و فيرمات . ألف جيمس كتاب بعنوان (فن التخمين The Art Of Conjecturing) . استغرق تأليف الكتاب زهاء عشرون عاماً ، و لم يطبع إلا عقب وفاته بثمانية أعوام . تضمّن الكتاب أربعة أجزاء ، الجزء الأول منه عبارة عن ملخص ما توصل إليه الهولندي هايجنز (1629 – 1695) من مفهوم هام في الاحتمالات و الرياضيات ألا و هو " التوقع الرياضي " و الذي أسماه حينها (قيمة الصدفة The value of the chance) . الجزء الثاني تضمّن التباديل و التوافيق . الجزء الثالث احتوى 24 مسألة متعلقة بألعاب الحظ و الصدفة التي كانت سائدة تلك الأيام . الجزء الأخير احتوى مسائل تطبيقية على الأخلاق و الاقتصاد و الشئون المدنية و على الرغم من أن جيمس لم يكمله إلا أنه يعتبر أهم إنجاز تطرق فيه لنقاش فلسفي مستفيض للمسائل المتعلقة بنظرية الاحتمالات و علاقة التوقع الرياضي بالأخلاق و توصل إلى أن الاحتمال ما هو إلا (مقياس لدرجة اليقين) . وضع برنوللي برهان للنظرية الشهيرة التي حملت اسمه و التي أطلق عليها بواسون اسم " قانون الأعداد الكبيرة " . أصبحت نظرية برنوللي حجر الزاوية في كثير من التطبيقات من رمي النرد و لعب الورق إلى الإحصاء الرياضي و نظرية الخطأ العشوائي و المسائل الديموغرافية . لقد

توصل برنوللي و بعد مخاض عسير في التجربة متعددة الرميات – مثل تجربة رمي قطعة النقود عدد نون من المرات – و التي فيها احتمال النجاح p و احتمال الفشل $q=1-p$ و جد برنوللي أن احتمال مشاهدة عدد r نجاح في عدد n رمية هو الحد الرائي من مفكوك ذات الحدين $(p+q)^n$ و الذي هو :-

$$P(n,r) = C(n,r)\, p^r q^{n-r}$$

دي موافر De moiver 1667-1745:

المعلم الآخر البارز في تطور نظرية الاحتمالات ما نشره دي موافر عام 1718 بعنوان (مذهب الصدفة doctrine of chance) و تضمّن حساب احتمالات الحوادث في الألعاب و التأمين مدى الحياة . مع ازدياد قيمة المضروب $n!=n(n-1)(n-2)\ldots3.2.1$

توصل دي موافر و جيمس سترلنغ إلى الصيغة :-

$$n! = \sqrt{2\pi n}\, n^n e^{-n}.$$

أعمال برنوللي و دي موافر أثارت اهتمام واسع و أدت إلى تطبيقات عريضة في تقدير الخطأ و دراسة النظم السياسية و التغيرات في التجمعات السكانية و الظواهر الاجتماعية (مثل متوسط طول العمر أو الزواج).

لابلاس P.S.Laplas 1749-1827 :-

هو الذي قاد نظرية الاحتمالات إلى أبعد من مجرد ألعاب الصدفة و الحظ ، قادها إلى آفاق علمية جادة و مفيدة . تفتقت عبقرية لابلاس في الاحتمالات و في نشوء نظرية الإحصاء الرياضي. تطرّق لابلاس في أبحاثه إلى عملية تحليل الخطأ المحتمل الوارد في الملاحظات . و كان من أروع ما كتب ما يعرف بقاعدة المربعات الصغرى لتصغير عدم اليقين في عدد من الملاحظات المستقلة . نشر سلسلة من الأبحاث جمعها في مؤلفه (نظرية التحليل الاحتمالية) عام 1812 . الجزء الأول يتعلق بحساب التفاضل و التكامل و الدوال المولدة . الجزء الثاني و

يتكون من نظرية الاحتمال المناسب و نظرية النهايات و نظرية الإحصاء الرياضي . تتأسس نظرية الصدفة من أن نختزل كل الحوادث من نفس النوع إلى رقم معيّن من الحالات متكافئة الفرص ، ثم وضع لابلاس التعريف الكلاسيكي للاحتمال : (احتمال ظهور حادثة ما هو النسبة بين عدد حالات ظهور تلك الحادثة إلى عدد ظهور كل الحالات الممكنة) و ذلك عندما نفتقر إلى اليقين الذي يحتم ظهور حادثة ما دون غيرها من الحوادث ، مع الأخذ بعين الاعتبار أن كل الحوادث متكافئة الفرص . صاغ لابلاس عملية جمع و ضرب الاحتمالات ، و تعرف العناصر الثلاثة الأولى من التالي ببديهيات الاحتمال :-

١ -قيمة الاحتمال دائماً لا سالبة

٢ -احتمال الحادثة المطلقة (أي تلك الصائبة منطقياً) قيمته الواحد ، ويتبع ذلك إن قيمة احتمال الحادثة المستحيلة (أي تلك الخاطئة منطقياً) صفراً .

٣ - جمع الاحتمالات :إذا كانت الحادثتان A , B متنافيتان (أي لا يمكن أن يحدثا معاً في نفس الوقت ، مثل حادثتا ظهور الطير أو النقش في عملية رمي قطعة النقود) فإن

$$P(AUB) = P(A) + P(B).$$

4- ضرب الاحتمالات : إذا كانت الحادثتان A , B مستقلتان (أي ظهور الحادثة A لا يؤثر في ظهور أو عدم ظهور الحادثة الأخرىB) فإن $P(A\cap B) = P(A).P(B)$

لابلاس و انطلاقاً من بيانات عينة إحصائية ، حسب النسبة بين عدد سكان بعض المقاطعات m إلى عدد المواليد السنوي n في تلك المقاطعات . و قدر عدد سكان فرنسا من الصيغة $P=N(m/n)$

حيث N هو إجمالي عدد المواليد السنوي للأمة . الخطأ الذي وقع فيه لابلاس أن اليقين الذي تتمتع به العلوم الطبيعية هو نفس اليقين الذي تتمتع به العلوم الإنسانية .

التوقع الرياضي : هو مجموع حاصل ضرب كل حادثة في احتمالها. مثلاً التوقع الرياضي عند رمي النرد هو :

2/7 = (6+5+4+3+2+1).(6/1)

يقال أن اللعبة عادلة إذا كان التوقع الرياضي لها صفر . التوقع الرياضي يعني متوسط ما يدفعه المراهن في عدد كثير من تكرار اللعبة . و بالتالي فإن مقدار 7/ 2الذي سيدفعه المراهن يعتبر مبلغ عادل لأنه إذا كرر الرهان لعدد كثير من المرات فإن متوسط ما سيربحه هو صفر . هناك مفارقة شهيرة في عملية التوقع الرياضي تعرف بمفارقة بطرسبيرغ St.Petersburg Paradox و كان أول ما صاغها نيكولاس بيرنوللي عام 1713ثم أعاد صياغتها دانيال بيرنوللي عام 1738على النحو التالي : لاعبان (أ) و (ب) اتفقا على أن يلعبا لعبة رمي قطعة نقود . تستمر اللعبة حتى تظهر أول صورة . اللاعب (ب) يعطي اللاعب (أ) قطعة نقود إذا ظهرت الصورة أولاً في أول رمية ، قطعتا نقود إذا ظهرت الصورة أولاً في ثاني رمية ، أربع قطع نقود إذا ظهرت الصورة أولاً في ثالث رمية ... و هكذا . ما هو المبلغ أي (التوقع الرياضي) الذي يجب أن يدفعه (أ) إلى (ب) – كرسوم – قبل بدء اللعبة حتى تكون اللعبة عادلة ؟ . لأنه لا يوجد حد منطقي أو نهاية لعدد النقش الذي يمكن أن يظهر قبل أن تظهر أول صورة يجب على (ب) أن يدفع $2^{ن-1}$ قطعة نقود إذا ما ظهر نقش في أول ن- 1 رمية قبل أن تظهر أول صورة . احتمال أن الصورة تظهر أول مرة في الرمية ن هي $(2/1)^{ن}$ ، و يصبح التوقع الرياضي ل (أ) هو :

$$...+ ^{ن}(2/1)^{ن-1}2 + ... + ^{3}(2/1).4+ ^{2}(2/1).2 + (2/1).1$$

= 2/1 + 2/1 + 2/1 + ... = لانهاية

يجب على اللاعب (أ) أن يدفع مقدماً مبلغ لا نهائي للاعب (ب) و هذا بالطبع يبدو سخيف . لكن مهما دفع اللاعب (أ) من مبلغ للاعب (ب) يظل هو الرابح في النهاية إذا تكررت اللعبة إلى عدد كبير جداً . على الرغم من (أ) سيربح آخر المطاف إلا أنه من المؤكد لن يجرؤ على دفع مبلغ خرافي مسبقاً كرسوم لدخول اللعبة . من الناحية العملية تم إجراء تجربة (1)[3] لرمي قطعة نقود و وجد في 2084 لعبة من رمي قطعة النقود أن 1061 أظهرت صورة من الرمية الأولى ، 494 في الرمية الثانية ، 232 في الرمية الثالثة ، 137 في الرابعة ، 56 في الخامسة ، 29 في السادسة ، 25 في السابعة ، 8 في الثامنة و 6 في التاسعة . وفق حساب التوقع الرياضي يجب على (أ) أن يدفع ل (ب) ما مجموعه 10057 قطعة نقود على ال 2084 لعبة كرسوم اشتراك (أو دخول) في اللعبة بما يجعل اللعبة عادلة نهاية المطاف . يكمن التناقض في مفارقة بطرسبيرغ بين الفكر الرياضي المجرد و بين الواقع التجريبي و الحس المشترك . المفارقة أثارت جدلاً واسعاً بين الرياضيين ردحاً من الزمان . العالم الرياضي دي لامبيرت ، لكي يتجاوز المفارقة افترض إمكانيتين : إمكانية طبيعية و إمكانية فوق طبيعية – ميتافيزيقية - . مثلاً أن الصورة لا تظهر إلا بعد ألف رمية فهذه إمكانية خارقة أي فوق طبيعية و هي بالطبع مستحيلة . و من ثم لا يوجد شخص موفق مطلقاً بل بالضرورة أن يتعثر حظه عند حد معين و تصبح المسألة في جوهرها مستحيلة . و تستمر مفارقة بطرسبورغ بلا حل و للأبد ! . التوقع الرياضي يعطي نتيجة لا معنى لها عندما يطبق على هذه المسألة . إلا إذا وضعنا قيد معين مثلاً بدل من أن يدفع (ب) متوالية من 1 ،2، ...، $2^{ن-1}$،... يدفع في المقابل متوالية من 1 ، ر ، $ر^2$ ، ... ، $ر^{ن-1}$ ، ...

(0> ر > 2) و التي مجموعها 1/(2- ر) هو التوقع الرياضي . واضح أن التوقع الرياضي هنا يصبح لانهائي عندما ر = 2

<u>الميكانيكا الإحصائي</u> يعتبر أول لبنة لنظرية فيزيائية تلعب فيها الاحتمالات دور أساسي . لقد أصبح من المعتقد منذ القرن السابع عشر أن الأنظمة المادية يمكن وصفها بعدد قليل من المعالم و التي يرتبط كلٌ منها مع الآخر بعلاقة تشبه القانون . هذه المعالم تعود إلى خواص المادة الهندسية ، الحركية و الحرارية . من هذه القوانين مثلاً قانون الغاز المثالي الذي يربط حاصل ضرب الضغط في الحجم بدرجة الحرارة لاحقاً أصبح الاتزان Equilibrium هو المفهوم الأساسي : الأنظمة بنفسها تغير قيم معالمها حتى تصل إلى الوضع الذي لا يلاحظ فيه أي تغير إضافي ، وضع الاتزان . مثلاً درجة الحرارة اللامنتظمة تتعدل عفوياً حتى تصبح منتظمة ، و نفس الشئ ينطبق على الكثافة . الانتروبي (القانون الثاني للديناميكا الحرارية) و هو معلم استحدثه R.Clausius ويعبر عن عجز النظام المعزول أن يصبح تلقائياً أكثر ترتيباً . أي لا ينخفض الانتروبي و بالتالي لا يتحقق الاتزان ، تناقض ! .

بدأت المفاهيم الاحتمالية في ولوج حلبة نظرية حركة الحرارة بواسطة ماكسويل و بولتزمان : يمكن حساب قيم اتزان الكميات بإدخال التوزيعات الاحتمالية على حالات الحركة المجهرية micro ، و استخدام متوسط الكميات المجهرية لتفسير و تعليل القيم المشاهدة للنظام الكبير macro برمته .

<u>التأويل الموضوعي للاحتمال Objective</u> : يتعاطى مع الاحتمال باعتباره تكرار للحادثة أو قياس لنزعة نتائج تجربة معينة . و السؤال المهم في التأويل الموضوعي : أنى للطبيعة أن يتمتع سلوكها بالطابع الاحتمالي ؟ .

التأويل الذاتي للاحتمال Subjective : يتعاطى مع الاحتمال باعتباره قياس لدرجة اليقين بما يتفق مع التفكير العقلاني و المنطقي و بما يتفق مع بديهيات الاحتمال .